FIGURE 1A

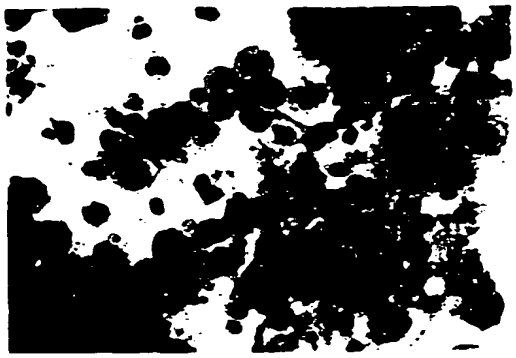


FIGURE 1B

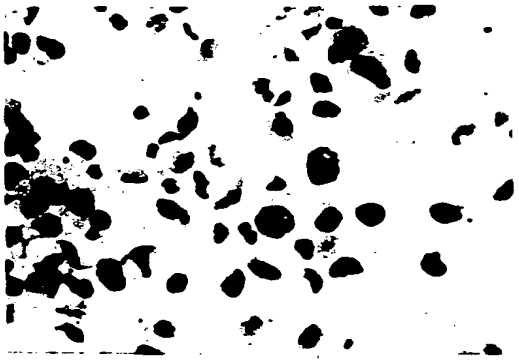


FIGURE 1C

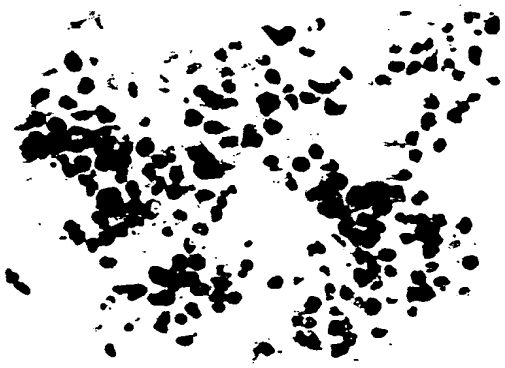
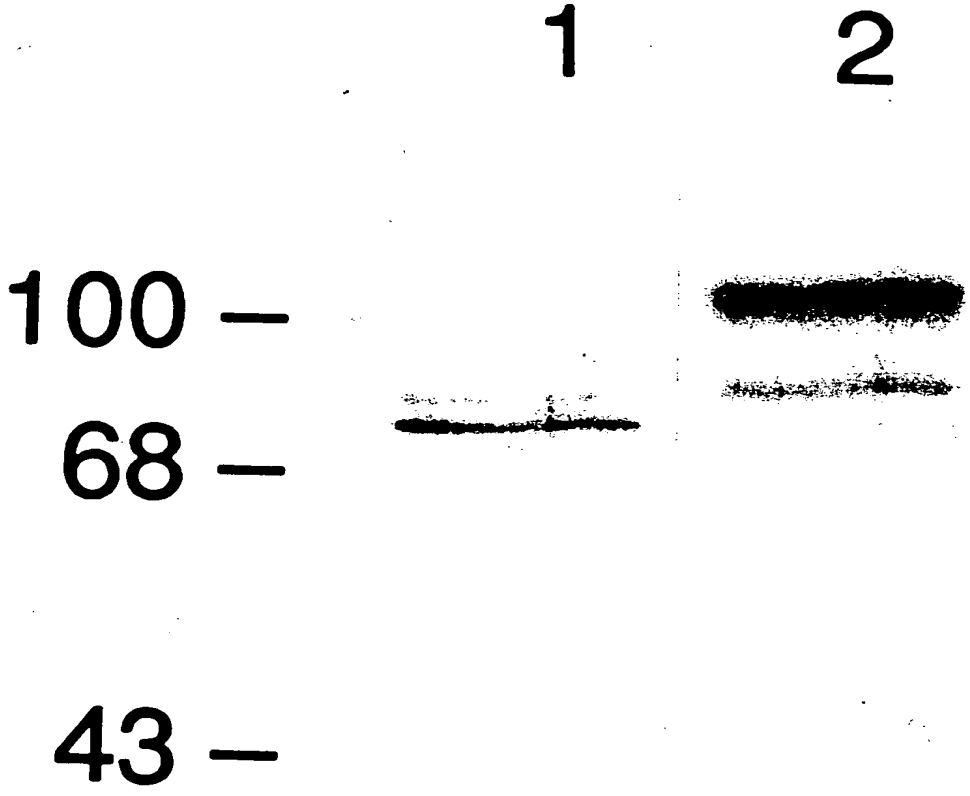


FIGURE 2



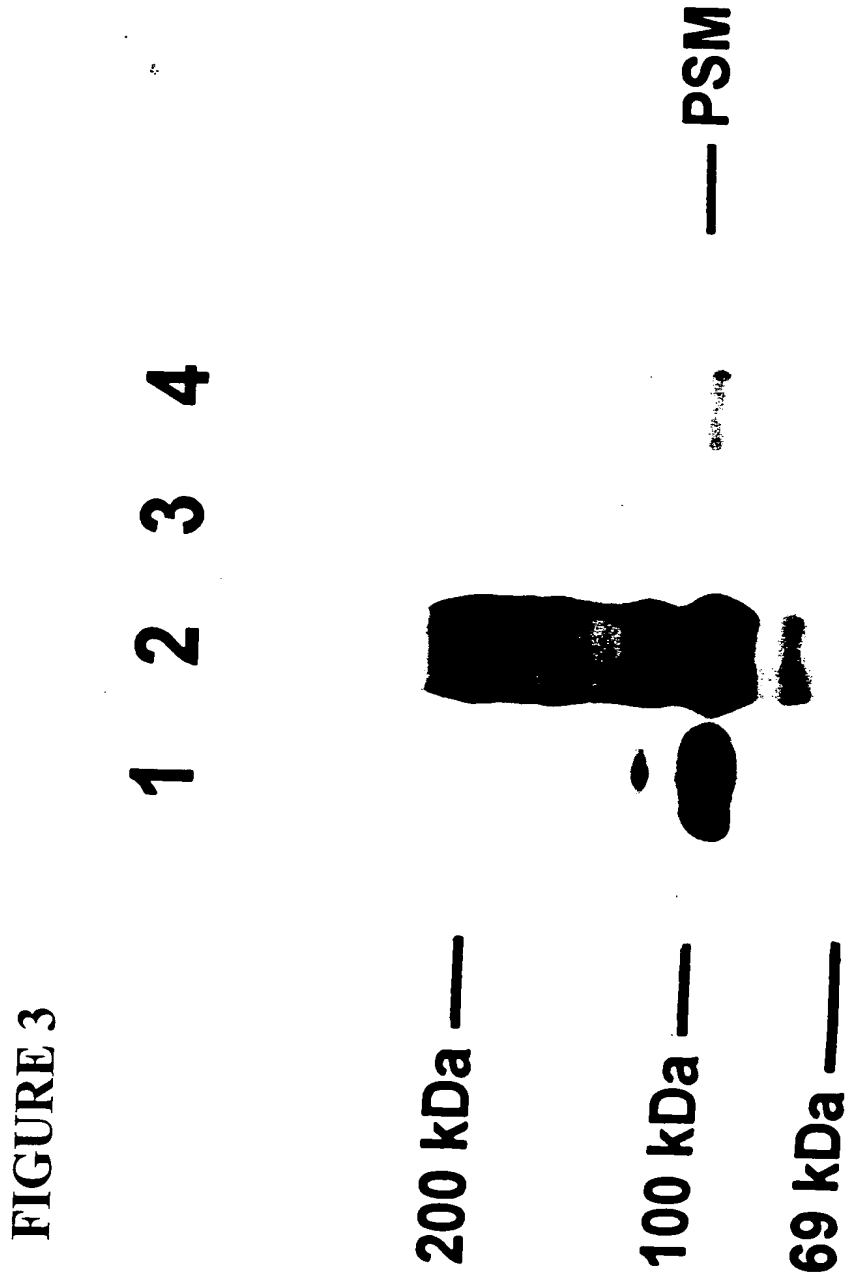


FIGURE 3

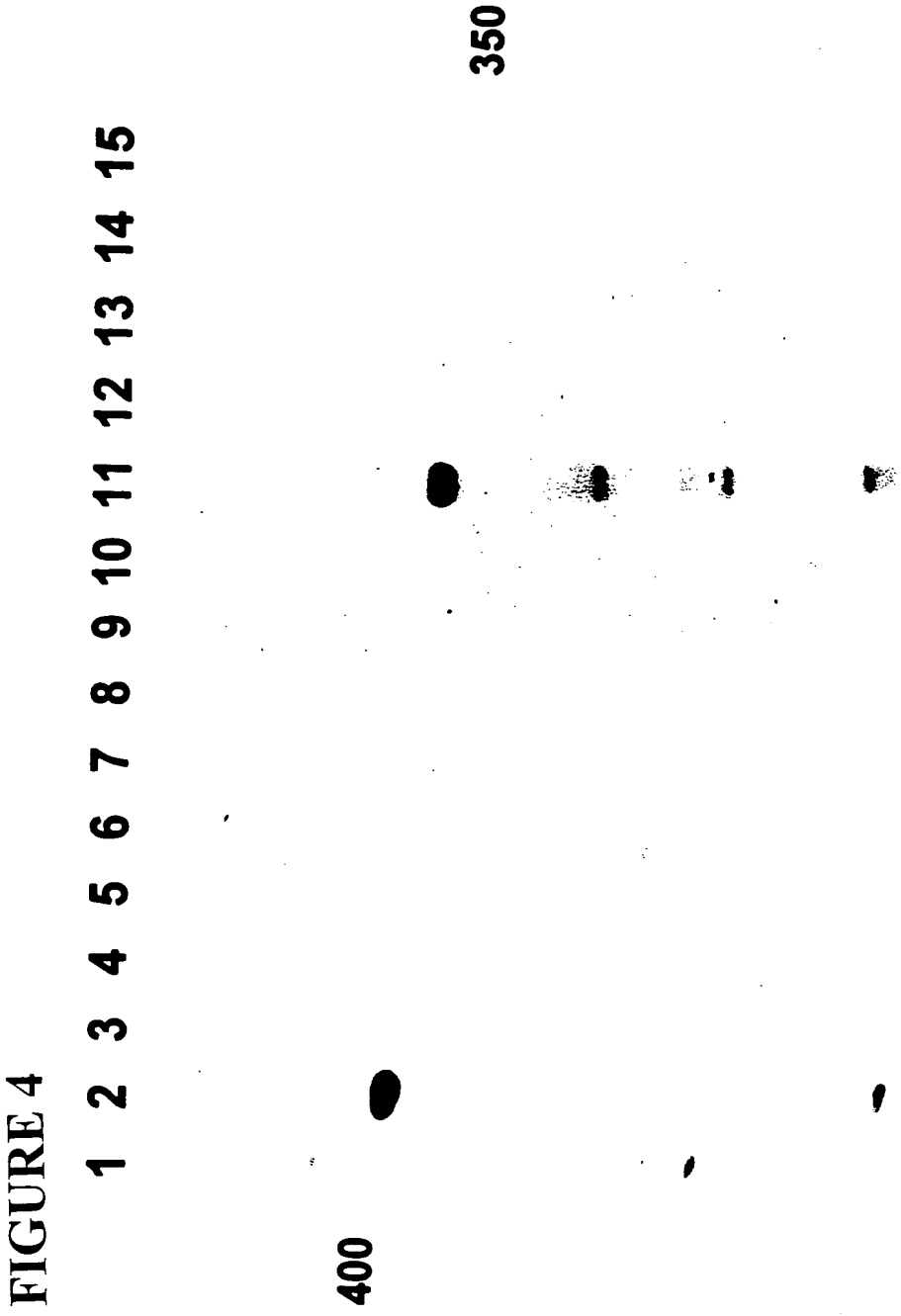


FIGURE 4

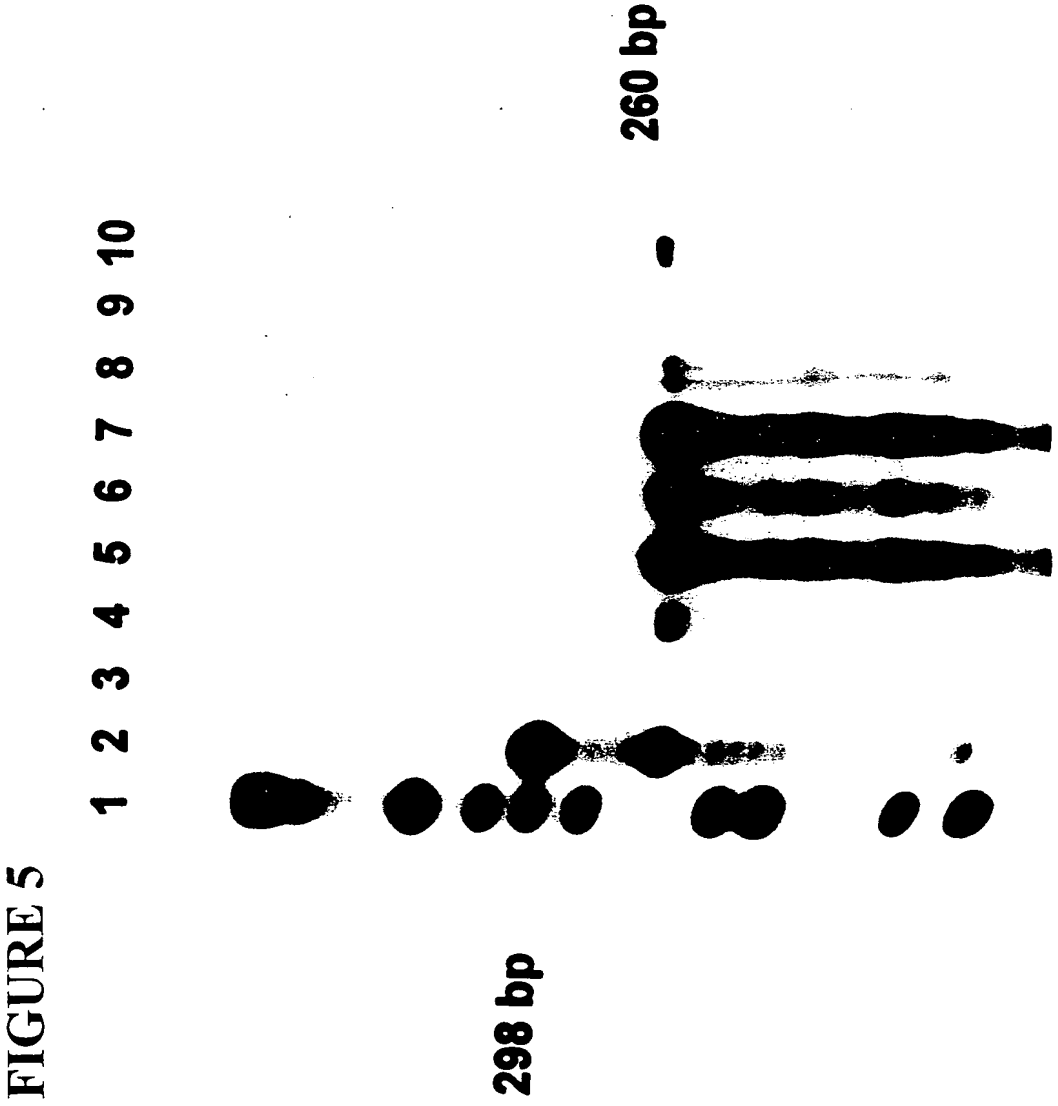


FIGURE 5

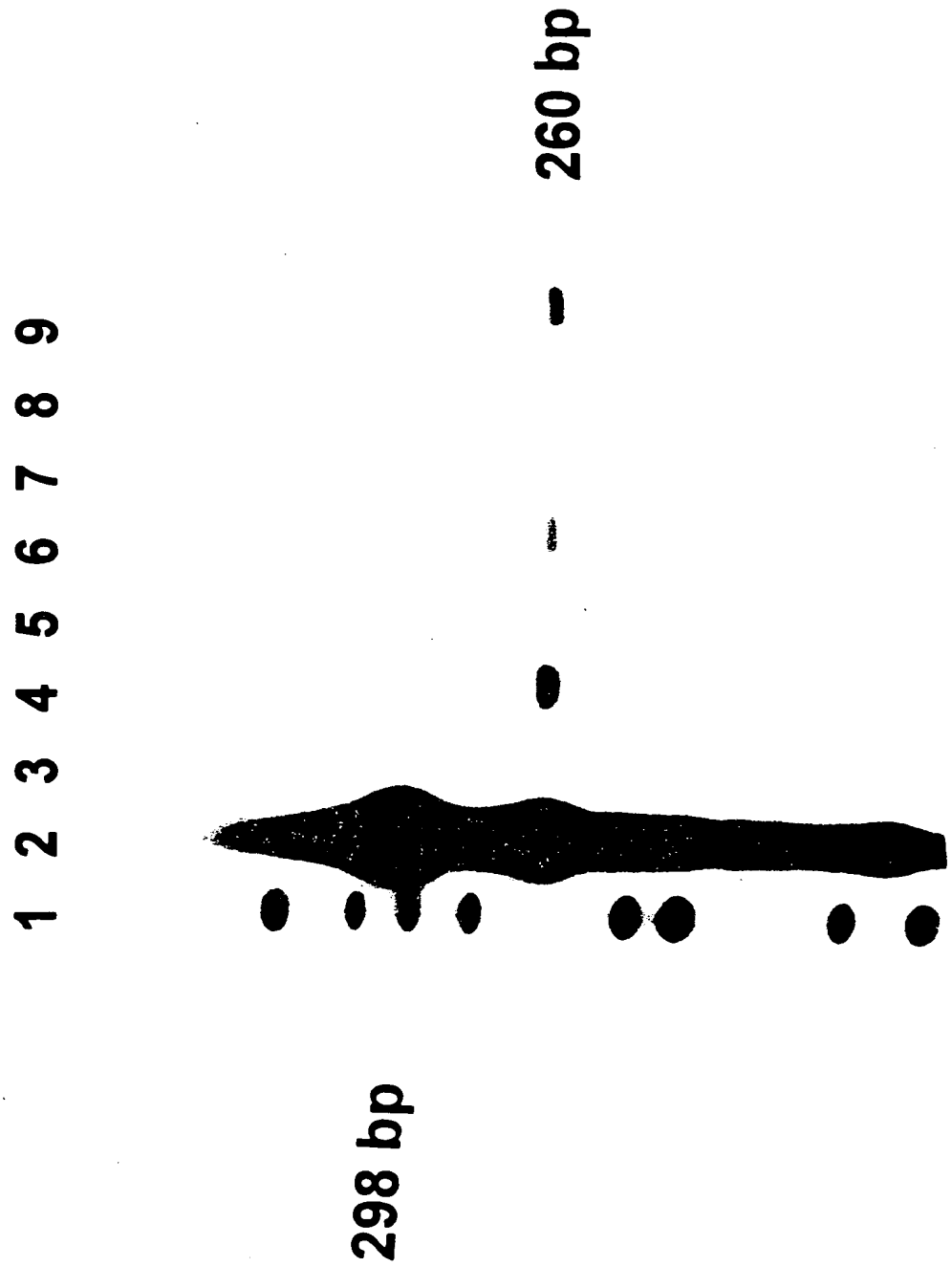


FIGURE 6

FIGURE 7

CELL LINE/TYPE	11p11.2-13 REGION	METASTATIC	PSM RNA DETECTED	PSM DNA DETECTED
LNCap			++	ND
HUMAN PROSTATE			++	ND
A9 (FIBROSARCOMA)	NO	NO	–	–
A9(11) (A9+HUM. 11)	YES	NO	–	REPEAT
AT6.1 (RAT PROSTATE)	NO	YES	–	–
AT6.1-11-c11	YES	NO	+	++
AT6.1-11-c12	NO	YES	–	–
R1564 (RAT MAMMARY)	NO	YES	–	–
R1564-11-c14	YES	YES	–	+
R1564-11-c15	YES	YES	–	REPEAT
R1564-11-c16	YES	YES	–	ND
R1564-11-c12	YES	YES	ND	+

FIGURE 8A

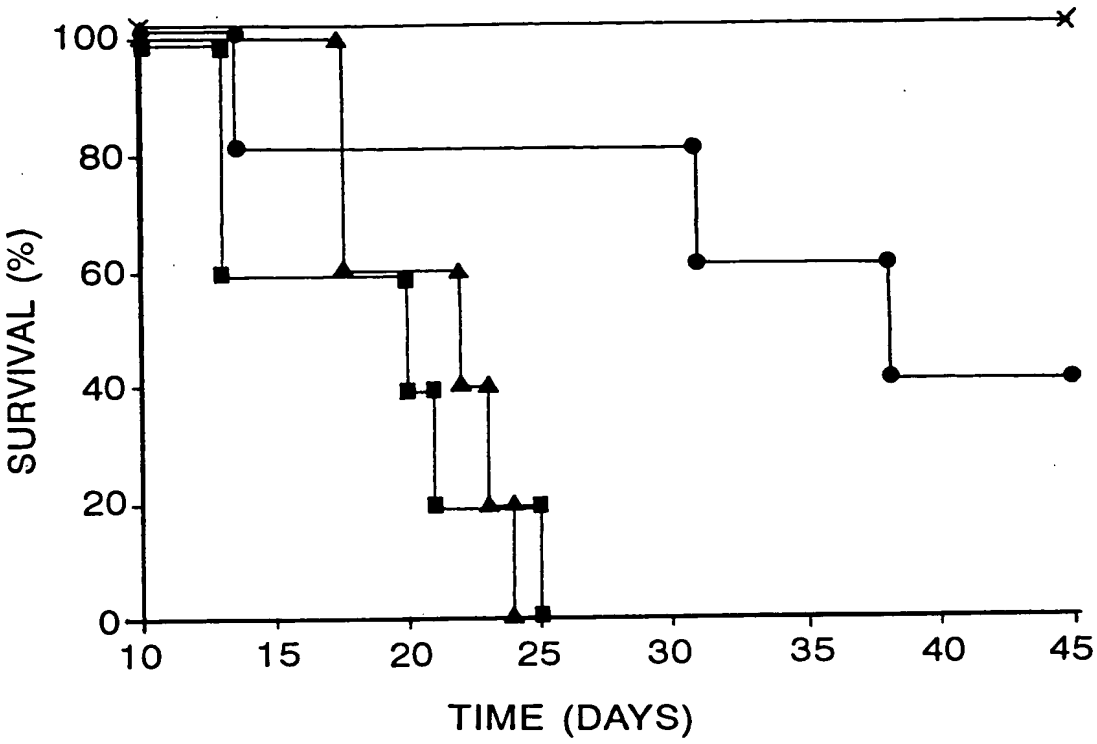


FIGURE 8B

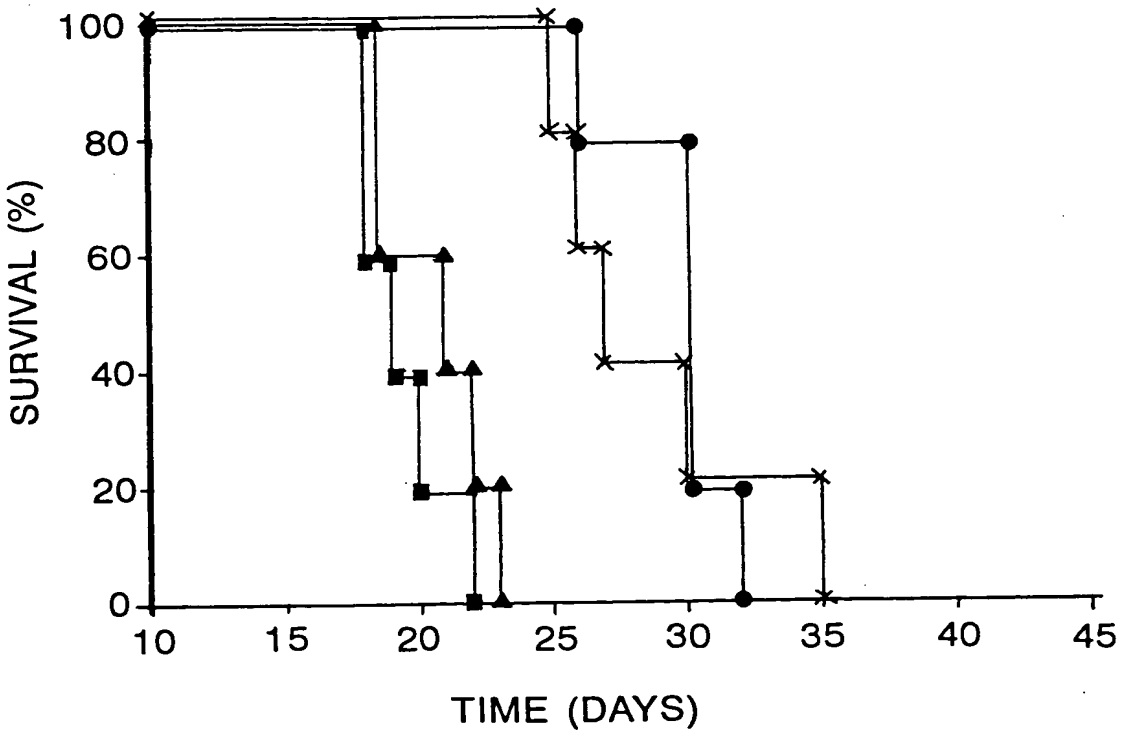


FIGURE 9A

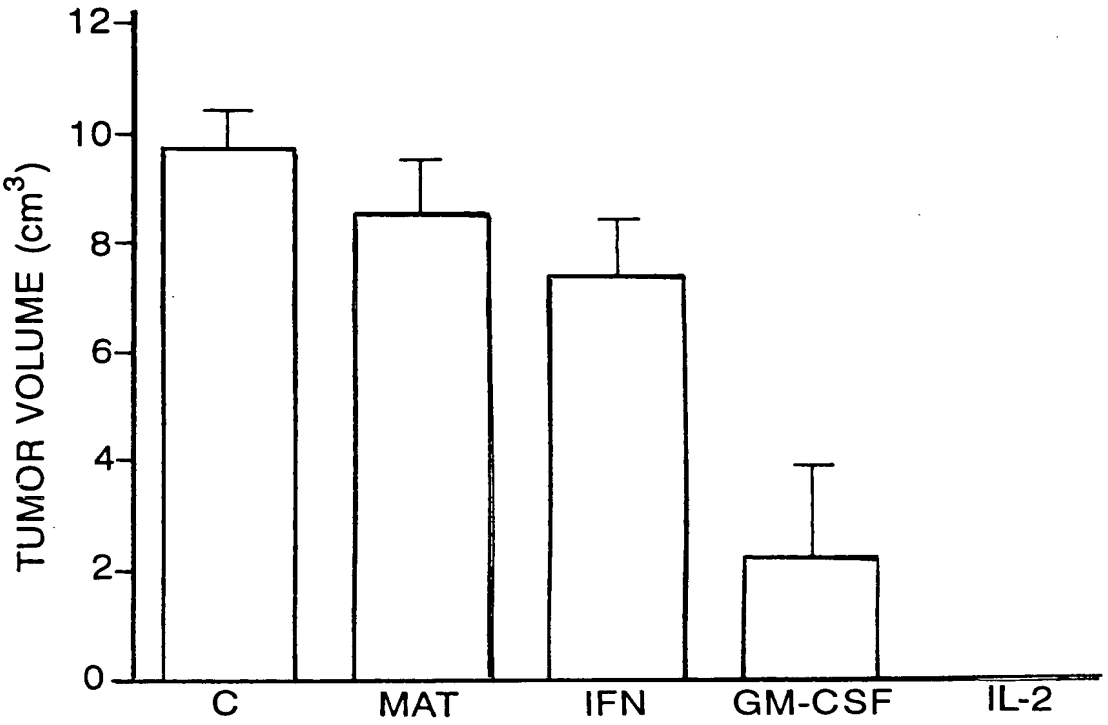
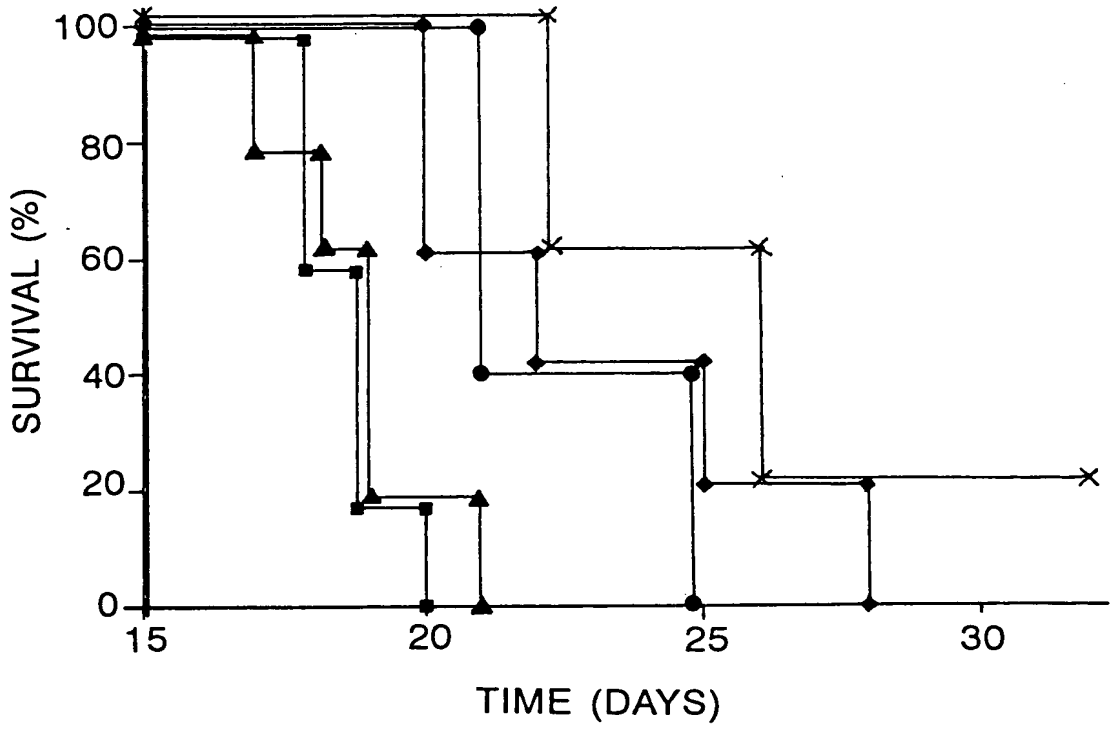


FIGURE 9B



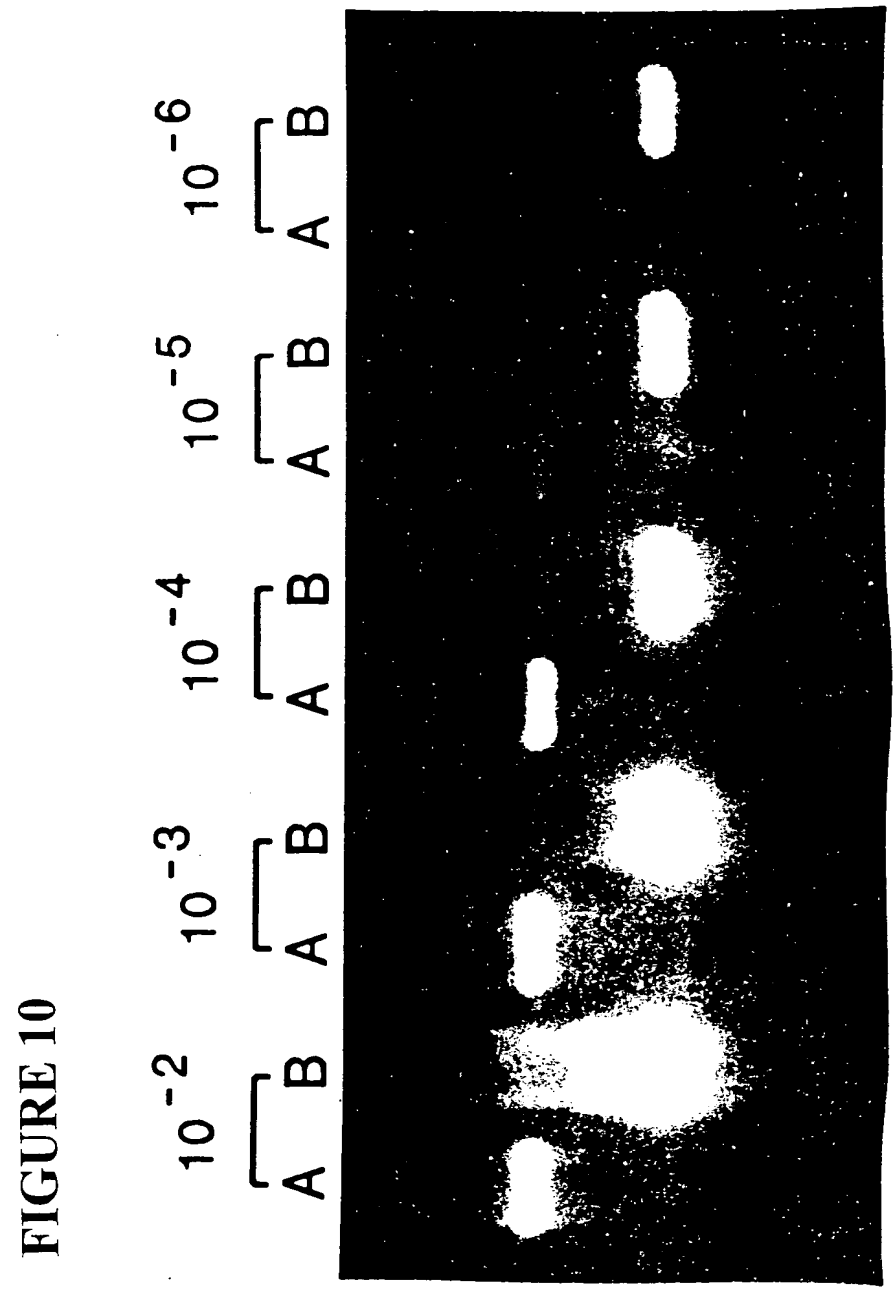


FIGURE 10

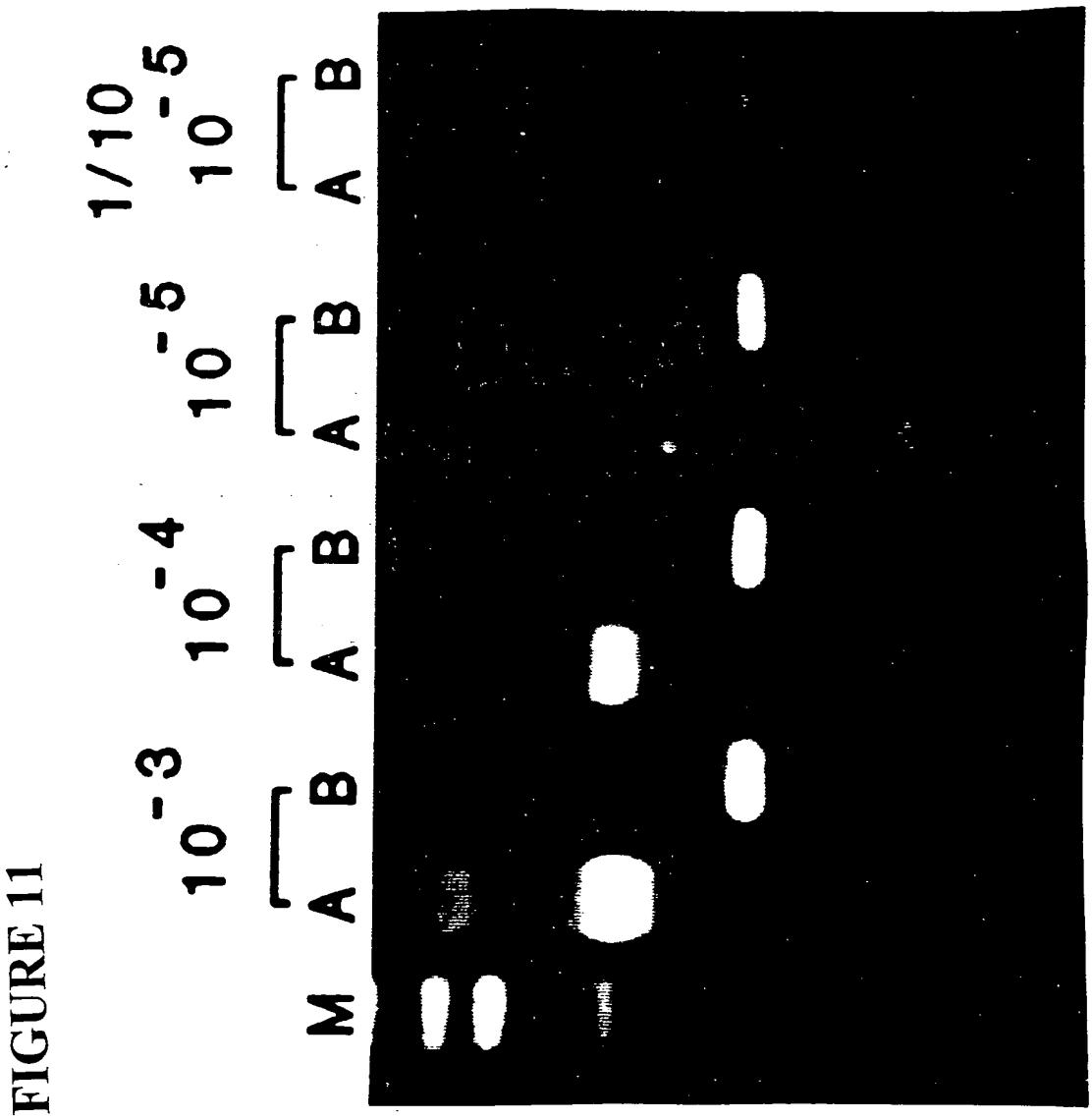


FIGURE 11

FIGURE 12

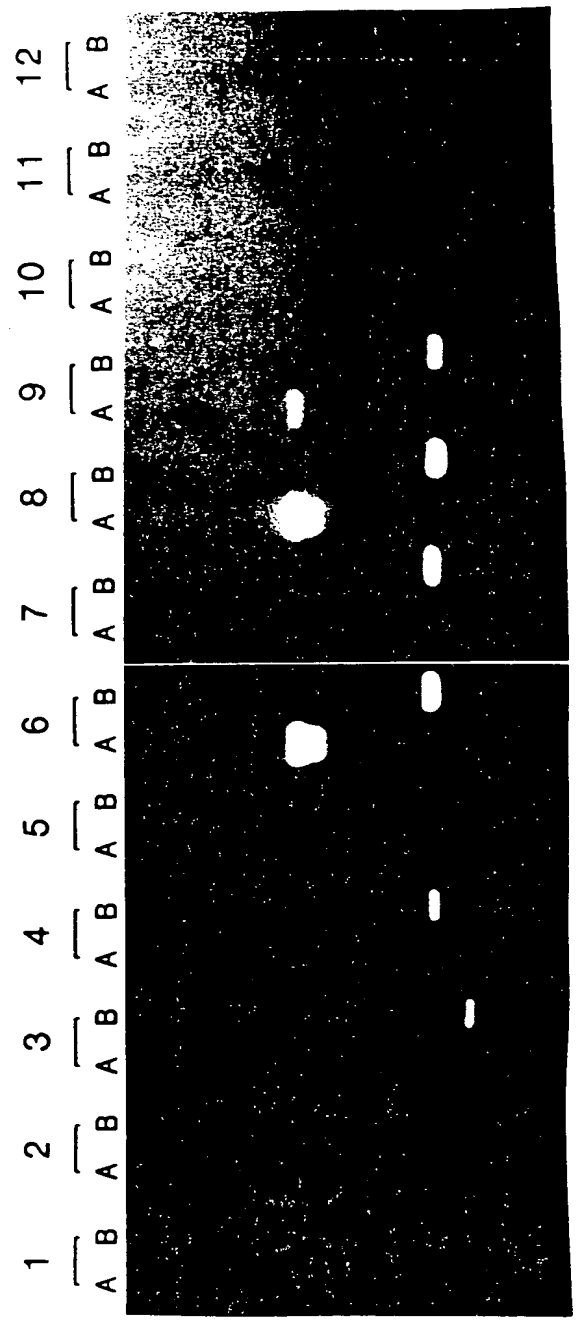


FIGURE 13

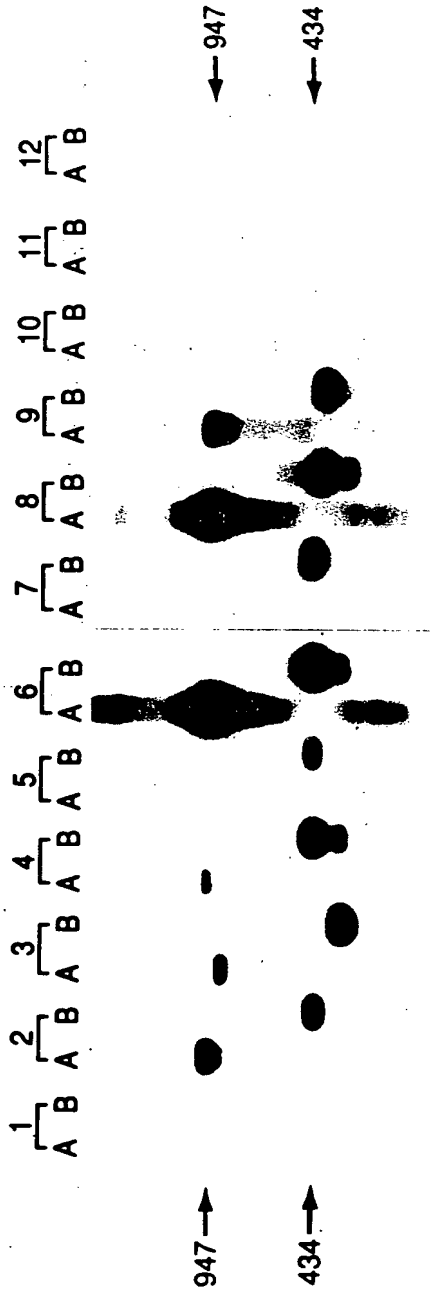

947
434
1 A B
2 A B
3 A B
4 A B
5 A B
6 A B
7 A B
8 A B
9 A B
10 A B
11 A B
12 A B
947
434

FIGURE 14

Patient	Stage	Treatment	PSA	PAP	PSA-PCR	PSM-PCR
1	T2NxMo	None	8.9	0.7	–	+
2	T2NoMo	RRP 7/93	6.1	–	–	+
3	T2CNoMo	PLND 5/93	4.5	0.1	–	+
4	T2BNoMo	RRP 3/92	NMA	0.4	–	+
5	T3NxMo	Proscar + Flutamide	51.3	1.0	–	+
6	Recur T3	I-125 1986	54.7	1.4	–	+
7	T3ANoMo	RRP 10/92	NMA	0.3	–	+
8	T3NxMo	XRT 1987	7.5	0.1	–	–
9	T3NxMo	Proscar + Flutamide	35.4	0.7	–	–
10	D2	S/P XRT Flutamide +Emcyt	311	4.5	+	+
11	D2	RRP 4/91 Lupron 10/92 Velban + Emcyt 12/92	1534	1.4	+	+
12	T2NoMo	RRP 8/91	NMA	0.5	–	+
13	T3NoMo	RRP 1/88 Lupron + Flutamide 5/92	0.1	0.3	–	–
14	D1	PLND 1989 XRT 1989	1.6	0.4	–	–
15	D1	Proscar + Flutamide	20.8	0.5	–	–
16	T2CNoMo	RRP 4/92	0.1	0.3	–	–

FIGURE 15A

```
          10         20         30         40         50         60
          |          |          |          |          |          |
  1 GCGCCTTAAA AAAAAAAAAC TTTCTTGGAA AATGTCCAGC TCTTGCTTAA ATATAAAAAT
    CGCGGAATTT TTTTTTTTTG AAAGAACCTT TTACAGGTCG AGAACGAATT TATATTTTTA

 61 GAAAGGAAGA AAGAGACTCT CCTCTCTCCA CTCCTATAAT TATGAGGAAC TTTTATTCAA
    CTTTCCTTCT TTCTCTGAGA GGAGAGAGGT GAGGATATTA ATACTCCTTG AAAATAAGTT

121 CTCTGAAATT CTATACAATC TCTACAATAC TCTACTGAAT AAAAGCAGAG CAGAAAAAGC
    GAGACTTTAA GATATGTTAG AGATGTTATG AGATGACTTA TTTTCGTCTC GTCTTTTTCG

181 TGCGCTTTTT TTCCATAGTC GGGAATGCTT GTCATCAGTG TAAATCACCA CCGCGCCCTT
    ACGCGAAAAA AAGGTATCAG CCCTTACGAA CAGTAGTCAC ATTTAGTGGT GGCGCGGGAA

241 TTTCCTAAAG AATATTATTG TTATTAATAA ACATGTAGGG TATTATCCTC CACTTACATT
    AAAGGATTTC TTATAATAAC AATAATTATT TGTACATCCC ATAATAGGAG GTGAATGTAA

301 ACAAAACCAT TTTTTAAAGC CGGGCGTGGT GGCTCACGCC TGTAATCCCA GCACTTTGGG
    TGTTTTGGTA AAAAATTTCG GCCCGCACCA CCGAGTGCGG ACATTAGGGT CGTGAAACCC

361 AGGCCCAGAC AGGCGGATCA CGAAGTCGAG AAATCGAGAC CATCCTGGCC AACATGGTGA
    TCCGGGTCTG TCCGCCTAGT GCTTCAGCTC TTTAGCTCTG GTAGGACCGG TTGTACCACT

421 AACCCCATCT CTACTAAAAA TACAAAAATT AGCTGGGCGT GGTGGCGGGC TCCTGTAGTC
    TTGGGGTAGA GATGATTTTT ATGTTTTTAA TCGACCCGCA CCACCGCCCG AGGACATCAG

481 CCAGCTACTC AGGAGGCTGA GGCAGGAGAA TCGCTTGAAC CGGGGAGGCG GAGGTTGCAG
    GGTCGATGAG TCCTCCGACT CCGTCCTCTT AGCGAACTTG GCCCCTCCGC CTCCAACGTC

541 TCAGCCAAGA TAGCGCCACT GCACTGGAGC CTGGTGACAG AGTGAGACTC CCTCAAGAAA
    AGTCGGTTCT ATCGCGGTGA CGTGACCTCG GACCACTGTC TCACTCTGAG GGAGTTCTTT

601 GAAAGGAAGG GAAGGGAAAG GGAAGGAAGG GGAGGGGAAG GGAGGGGAGG GGAGGGGAGG
    CTTTCCTTCC CTTCCCTTTC CCTTCCTTCC CCTCCCCTTC CCTCCCCTCC CCTCCCCTCC

661 AAAGAAAAGA ATACTGGAAC TTGTTGAAGG CAGAGACTTT ATTTTCATAT CCCGGCTATG
    TTTCTTTTCT TATGACCTTG AACAACTTCC GTCTCTGAAA TAAAAGTATA GGGCCGATAC

721 TCTGGCTACT GTCTTACGTA ATAGATATAA AATCAATCTT GGTTGGATTA ACCAGAAGAA
    AGACCGATGA CAGAATGCAT TATCTATATT TTAGTTAGAA CCAACCTAAT TGGTCTTCTT
```

FIGURE 15B

```
 781 TGAGAAGATA TATTCTGGTA AGTTGAATAC TTAGCACCCA GGGGTAATCA GCTTGGACAG
     ACTCTTCTAT ATAAGACCAT TCAACTTATG AATCGTGGGT CCCCATTAGT CGAACCTGTC

 841 GACCAGGTCC AAAGACTGTT AAGAGTCTTC TGACTCCAAA CTCAGTGCTC CCTCCAGTGC
     CTGGTCCAGG TTTCTGACAA TTCTCAGAAG ACTGAGGTTT GAGTCACGAG GGAGGTCACG

 901 CACAAGCAAA CTCCATAAAG GTATCCTGTG CTGAATAGAG ACTGTAGAGT GGTACAAAGT
     GTGTTCGTTT GAGGTATTTC CATAGGACAC GACTTATCTC TGACATCTCA CCATGTTTCA

 961 AAGACAGACA TTATATTAAG TCTTAGCTTT GTGACTTCGA ATGACTTACC TAATCTAGCT
     TTCTGTCTGT AATATAATTC AGAATCGAAA CACTGAAGCT TACTGAATGG ATTAGATCGA

1021 AAATTTCAGT TTTACCATGT GTAAATCAGG AAGAGTAATA GAACAAACCT TGAAGGGTCC
     TTTAAAGTCA AAATGGTACA CATTTAGTCC TTCTCATTAT CTTGTTTGGA ACTTCCCAGG

1081 CAATGGTGAT TAAATGAGGT GATGTACATA ACATGCATCA CTCATAATAA GTGCTCTTTA
     GTTACCACTA ATTTACTCCA CTACATGTAT TGTACGTAGT GAGTATTATT CACGAGAAAT

1141 AATATTAGTC ACTATTATTA GCCATCTCTG ATTAGATTTG ACAATAGGAA CATTAGGAAA
     TTATAATCAG TGATAATAAT CGGTAGAGAC TAATCTAAAC TGTTATCCTT GTAATCCTTT

1201 GATATAGTAC ATTCAGGATT TTGTTAGAAA GAGATGAAGA AATTCCCTTC CTTCCTGCCC
     CTATATCATG TAAGTCCTAA AACAATCTTT CTCTACTTCT TTAAGGGAAG GAAGGACGGG

1261 TAGGTCATCT AGGAGTTGTC ATGGTTCATT GTTGACAAAT TAATTTTCCC AAATTTTTCA
     ATCCAGTAGA TCCTCAACAG TACCAAGTAA CAACTGTTTA ATTAAAAGGG TTTAAAAAGT

1321 CTTTGCTCAG AAAGTCTACA TCGAAGCACC CAAGACTGTA CAATCTAGTC CATCTTTTTC
     GAAACGAGTC TTTCAGATGT AGCTTCGTGG GTTCTGACAT GTTAGATCAG GTAGAAAAAG

1381 CACTTAACTC ATACTGTGCT CTCCCTTTCT CAAAGCAAAC TGTTTGCTAT TCCTTGAATA
     GTGAATTGAG TATGACACGA GAGGGAAAGA GTTTCGTTTG ACAAACGATA AGGAACTTAT

1441 CACTCTGAGT TTTCTGCCTT TGCCTACTCA GCTGGCCCAT GGCCCCTAAT GTTTCTTCTC
     GTGAGACTCA AAAGACGGAA ACGGATGAGT CGACCGGGTA CCGGGGATTA CAAAGAAGAG

1501 ATCTCCACTG GGTCAAATCC TACCTGTACC TTATGGTTCT GTTAAAAGCA GTGCTTCCAT
     TAGAGGTGAC CCAGTTTAGG ATGGACATGG AATACCAAGA CAATTTTCGT CACGAAGGTA

1561 AAAGTACTCC TAGCAAATGC ACGGCCTCTC TCACGGATTA TAAGAACACA GTTTATTTTA
     TTTCATGAGG ATCGTTTACG TGCCGGAGAG AGTGCCTAAT ATTCTTGTGT CAAATAAAAT

1621 TAAAGCATGT AGCTATTCTC TCCCTCGAAA TACGATTATT ATTATTAAGA ATTTATAGCA
     ATTTCGTACA TCGATAAGAG AGGGAGCTTT ATGCTAATAA TAATAATTCT TAAATATCGT

1681 GGGATATAAT TTTGTATGAT GATTCTTCTG GTTAATCCAA CCAAGATTGA TTTTATATCT
     CCCTATATTA AAACATACTA CTAAGAAGAC CAATTAGGTT GGTTCTAACT AAAATATAGA

1741 ATTACGTAAG ACAGTAGCCA GACATAGCCG GGATATGAAA ATAAAGTCTC TGCCTTCAAC
     TAATGCATTC TGTCATCGGT CTGTATCGGC CCTATACTTT TATTTCAGAG ACGGAAGTTG

1801 AAGTTCCAGT ATTCTTTTCT TTCCTCCCCT CCCCTCCCCT CCCTTCCCCT CCCCTTCCTT
     TTCAAGGTCA TAAGAAAAGA AAGGAGGGGA GGGGAGGGGA GGGAAGGGGA GGGGAAGGAA

1861 CCCTTTCCCT TCCCTTCCTT TCTTTCTTGA GGGAGTCTCA CTCTGTCACC AGGCTCCAGT
     GGGAAAGGGA AGGGAAGGAA AGAAAGAACT CCCTCAGAGT GAGACAGTGG TCCGAGGTCA
```

FIGURE 15C

```
1921 GCAGTGGCGC TATCTTGGCT GACTGCAACC TCCGCCTCCC CGGTTCAAGC GATTCTCCTG
     CGTCACCGCG ATAGAACCGA CTGACGTTGG AGGCGGAGGG GCCAAGTTCG CTAAGAGGAC

1981 CCTCAGCCTC CTGAGTAGCT GGGACTACAG GAGCCCGCCA CCACGCCCAG CTAATTTTTG
     GGAGTCGGAG GACTCATCGA CCCTGATGTC CTCGGGCGGT GGTGCGGGTC GATTAAAAAC

2041 TATTTTTAGT AGAGATGGGG TTTCACCATG TTGGCCAGGA TGGTCTCGAT TTCTCGACTT
     ATAAAAATCA TCTCTACCCC AAAGTGGTAC AACCGGTCCT ACCAGAGCTA AAGAGCTGAA

2101 CGTGATCCGC CTGTCTGGGC CTCCCAAAGT GCTGGGATTA CAGGCGTGAG CCACCACGCC
     GCACTAGGCG GACAGACCCG GAGGGTTTCA CGACCCTAAT GTCCGCACTC GGTGGTGCGG

2161 CGGCTTTAAA AAATGGTTTT GTAATGTAAG TGGAGGATAA TACCCTACAT GTTTATTAAT
     GCCGAAATTT TTTACCAAAA CATTACATTC ACCTCCTATT ATGGGATGTA CAAATAATTA

2221 AACAATAATA TTCTTTAGGA AAAAGGGCGC GGTGGTGATT TACACTGATG ACAAGCATTC
     TTGTTATTAT AAGAAATCCT TTTTCCCGCG CCACCACTAA ATGTGACTAC TGTTCGTAAG

2281 CCGACTATGG AAAAAAAGCG CAGCTTTTTC TGCTCTGCTT TTATTCAGTA GAGTATTGTA
     GGCTGATACC TTTTTTTCGC GTCGAAAAAG ACGAGACGAA AATAAGTCAT CTCATAACAT

2341 GAGATTGTAT AGAATTTCAG AGTTGAATAA AAGTTCCTCA TAATTATAGG AGTGGAGAGA
     CTCTAACATA TCTTAAAGTC TCAACTTATT TTCAAGGAGT ATTAATATCC TCACCTCTCT

2401 GGAGAGTCTC TTTCTTCCTT TCATTTTTAT ATTTAAGCAA GAGCTGGACA TTTTCCAAGA
     CCTCTCAGAG AAAGAAGGAA AGTAAAAATA TAAATTCGTT CTCGACCTGT AAAAGGTTCT

2461 AAGTTTTTTT TTTTTAAGGC GCCTCTCAAA AGGGGCCGGA TTTCCTTCTC CTGGAGGCAG
     TTCAAAAAAA AAAAATTCCG CGGAGAGTTT TCCCCGGCCT AAAGGAAGAG GACCTCCGTC

2521 ATGTTGCCTC TCTCTCTCGC TCGGATTGGT TCAGTGCACT CTAGAAACAC TGCTGTGGTG
     TACAACGGAG AGAGAGAGCG AGCCTAACCA AGTCACGTGA GATCTTTGTG ACGACACCAC

2581 GAGAAACTGG ACCCCAGGTC TGGAGCGAAT TCCAGCCTGC AGGGCTGATA AGCGAGGCAT
     CTCTTTGACC TGGGGTCCAG ACCTCGCTTA AGGTCGGACG TCCCGACTAT TCGCTCCGTA

2641 TAGTGAGATT GAGAGAGACT TTACCCCGCC GTGGTGGTTG GAGGGCGCGC AGTAGAGCAG
     ATCACTCTAA CTCTCTCTGA AATGGGGCGG CACCACCAAC CTCCCGCGCG TCATCTCGTC

2701 CAGCACAGGC CCGGGTCCCG GGAGGCCGGC TCTGCTCGCG CCGAGATGTG GAATCTCCTT
     GTCGTGTCCG CGCCCAGGGC CCTCCGGCCG AGACGAGCGC GGCTCTACAC CTTAGAGGAA

2761 CACGAAACCG ACTCGGCTGT GGCCACCGCC CGCCGCCCGC GCTGGCTGTG CGCTGGGGCG
     GTGCTTTGGC TGAGCCGACA CCGGTGGCGC GCGGCGGGCG CGACCGACAC GCGACCCCGC

2821 CTGGTGCTGG CGGGTGGCTT CTTTCTCCTC GGCTTCCTCT TCGGTAGGGG GGCGCCTCGC
     GACCACGACC GCCCACCGAA GAAAGAGGAG CCGAAGGAGA AGCCATCCCC CCGCGGAGCG

2881 GGAGCAAACC TCGGAGTCTT CCCCGTGGTG CCGCGGTGCT GGGACTCGCG GGTCAGCTGC
     CCTCGTTTGG AGCCTCAGAA GGGGCACCAC GGCGCCACGA CCCTGAGCGC CCAGTCGACG

2941 CGAGTGGGAT CCTGTTGCTG GTCTTCCCCA GGGGCGGCGA TTAGGGTCGG GGTAATGTGG
     GCTCACCCTA GGACAACGAC CAGAAGGGGT CCCCGCCGCT AATCCCAGCC CCATTACACC

3001 GGTGAGCACC CCTCGAG
     CCACTCGTGG GGAGCTC
```

FIGURE 15D


2401 GGAGAGTCTC TTTCTTCCTT TCATTTTTAT ATTTAAGCAA GAGCTGGACA TTTTCCAAGA
CCTCTCAGAG AAAGAAGGAA AGTAAAAATA TAAATTCGTT CTCGACCTGT AAAAGGTTCT
2461 AAGTTTTTTT TTTTTAAGGC GCCTCTCAAA AGGGGCCGGA TTTCCTTCTC CTGGAGGCAG
TTCAAAAAAA AAAAATTCCG CGGAGAGTTT TCCCCGGCCT AAAGGAAGAG GACCTCCGTC
2521 ATGTTGCCTC TCTCTCTCGC TCGGATTGGT TCAGTGCACT CTAGAAACAC TGCTGTGGTG
TACAACGGAG AGAGAGAGCG AGCCTAACCA AGTCACGTGA GATCTTTGTG ACGACACCAC
2581 GAGAAACTGG ACCCCAGGTC TGGAGCGAAT TCCAGCCTGC AGGGCTGATA AGCGAGGCAT
CTCTTTGACC TGGGGTCCAG ACCTCGCTTA AGGTCGGACG TCCCGACTAT TCGCTCCGTA
2641 TAGTGAGATT GAGAGAGACT TTACCCCGCC GTGGTGGTTG GAGGGCGCGC AGTAGAGCAG
ATCACTCTAA CTCTCTCTGA AATGGGGCGG CACCACCAAC CTCCCGCGCG TCATCTCGTC
2701 CAGCACAGGC GCGGGTCCCG GGAGGCCGGC TCTGCTCGCG CCGAGATGTG GAATCTCCTT
GTCGTGTCCG CGCCCAGGGC CCTCCGGCCG AGACGAGCGC GGCTCTACAC CTTAGAGGAA
2761 CACGAAACCG ACTCGGCTGT GGCCACCGCG CGCCGCCCGC GCTGGCTGTG CGCTGGGGCG
GTGCTTTGGC TGAGCCGACA CCGGTGGCGC GCGGCGGGCG CGACCGACAC GCGACCCCGC
2821 CTGGTGCTGG CGGGTGGCTT CTTTCTCCTC GGCTTCCTCT TCGGTAGGGG GGCGCCTCGC
GACCACGACC GCCCACCGAA GAAAGAGGAG CCGAAGGAGA AGCCATCCCC CCGCGGAGCG
2881 GGAGCAAACC TCGGAGTCTT CCCCGTGGTG CCGCGGTGCT GGGACTCGCG GGTCAGCTGC
CCTCGTTTGG AGCCTCAGAA GGGGCACCAC GGCGCCACGA CCCTGAGCGC CCAGTCGACG
2941 CGAGTGGGAT CCTGTTGCTG GTCTTCCCCA GGGGCGGCGA TTAGGGTCGG GGTAATGTGG
GCTCACCCTA GGACAACGAC CAGAAGGGGT CCCCGCCGCT AATCCCAGCC CCATTACACC
3001 GGTGAGCACC CCTCGAG
CCACTCGTGG GGAGCTC

FIGURE 16

Potential binding sites on the PSM promoter*

Site	Seq	**Location	#nt matched	
AP1	TKAGTCA	-1145	7/7	
E2-RS	ACCNNNNNNGGT	-1940	12/12	
		-1951	12/12	
GHF	NNNTAAATNNN	-580	11/11	
		-753	11/11	
		-1340	11/11	
		-1882	11/11	
		-1930	11/11	
		-1979	11/11	
		-2001	11/11	
		-2334	11/11	
		-2374	11/11	
		-2591	11/11	
		-2620	11/11	
		-2686	11/11	
JVC repeat	GGGNGGRR	-1165	8/8	
		-1175	8/8	
		-1180	8/8	
		-1185	8/8	
		-1190	8/8	
NFkB	GGGRHTYYHC	-961	10/10	
uteroglobi	RYYWSGTG	-250	8/8	
		-921	8/8	
		-1104	8/8	
IFN	AAWAANGAAAGGR	590	13/13	Cell 41:509 (1985)

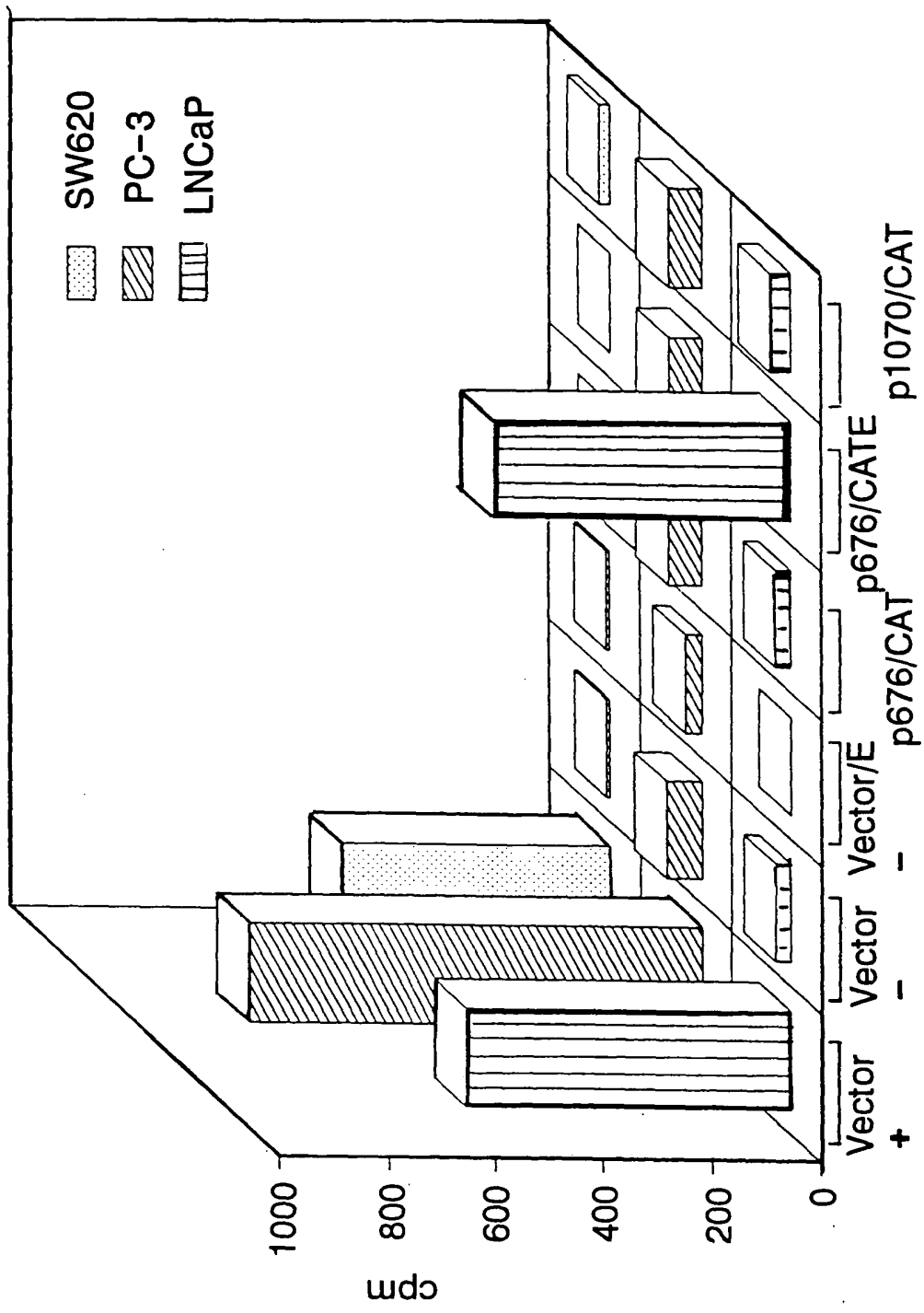
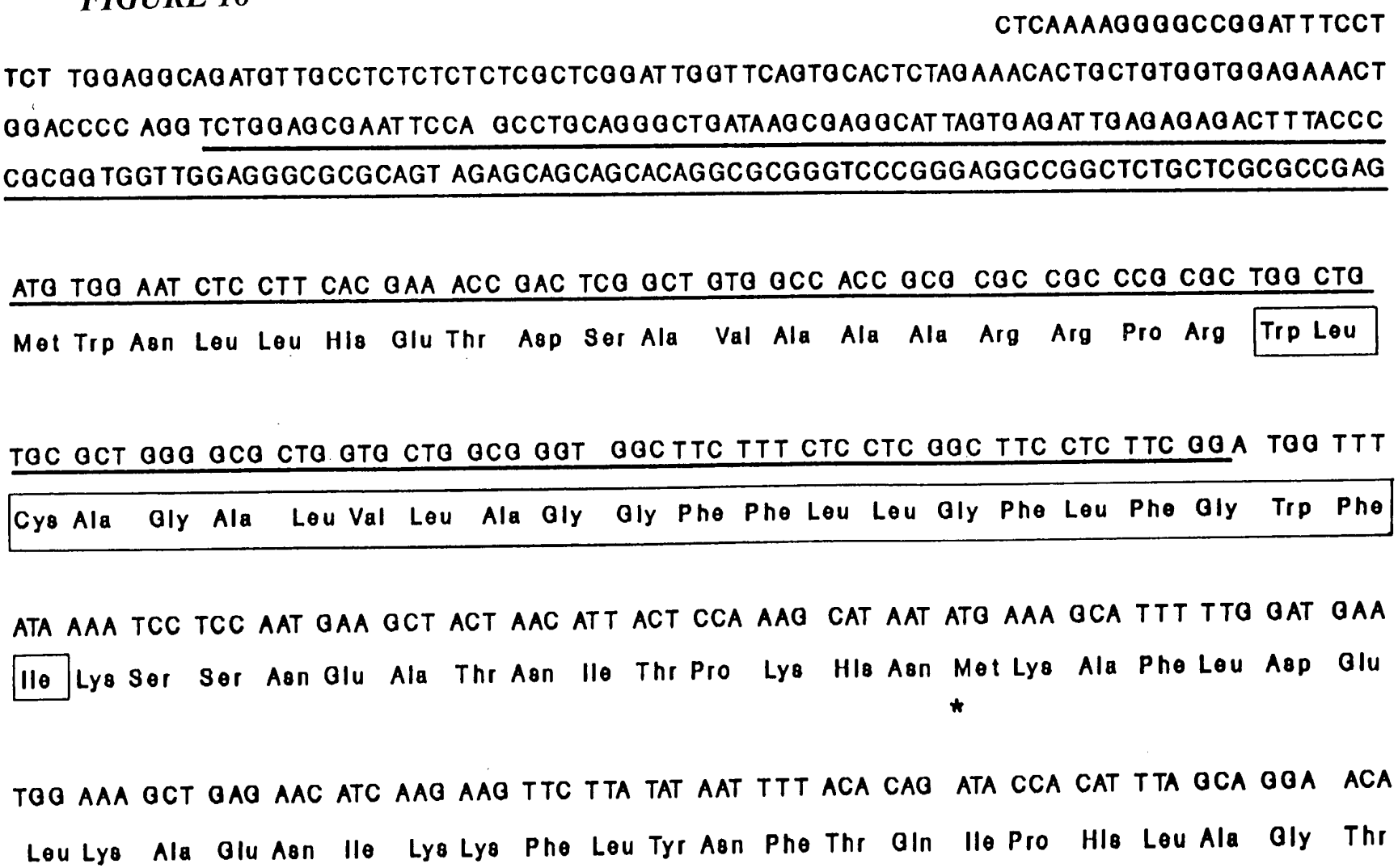

SW620
PC-3
LNCaP
cpm
1000
800
600
400
200
0
+
-
-
Vector
Vector
Vector/E
p676/CAT
p676/CATE
p1070/CAT

FIGURE 17

FIGURE 18


CTCAAAAGGGGCCGGATTTCCT
TCT TGGAGGCAGATGTTGCCTCTCTCTCTCGCTCGGATTGGTTCAGTGCACTCTAGAAACACTGCTGTGGTGGAGAAACT
GGACCCC AGG TCTGGAGCGAATTCCA GCCTGCAGGGCTGATAAGCGAGGCATTAGTGAGATTGAGAGAGACTTTACCC
CGCGGTGGTTGGAGGGCGCGCAGT AGAGCAGCAGCACAGGCGCGGGTCCCGGGAGGCCGGCTCTGCTCGCGCCGAG
ATG TGG AAT CTC CTT CAC GAA ACC GAC TCG GCT GTG GCC ACC GCG CGC CGC CCG CGC TGG CTG
Met Trp Asn Leu Leu His Glu Thr Asp Ser Ala Val Ala Ala Ala Arg Arg Pro Arg Trp Leu
TGC GCT GGG GCG CTG GTG CTG GCG GGT GGC TTC TTT CTC CTC GGC TTC CTC TTC GGA TGG TTT
Cys Ala Gly Ala Leu Val Leu Ala Gly Gly Phe Phe Leu Leu Gly Phe Leu Phe Gly Trp Phe
ATA AAA TCC TCC AAT GAA GCT ACT AAC ATT ACT CCA AAG CAT AAT ATG AAA GCA TTT TTG GAT GAA
Ile Lys Ser Ser Asn Glu Ala Thr Asn Ile Thr Pro Lys His Asn Met Lys Ala Phe Leu Asp Glu
*
TGG AAA GCT GAG AAC ATC AAG AAG TTC TTA TAT AAT TTT ACA CAG ATA CCA CAT TTA GCA GGA ACA
Leu Lys Ala Glu Asn Ile Lys Lys Phe Leu Tyr Asn Phe Thr Gln Ile Pro His Leu Ala Gly Thr

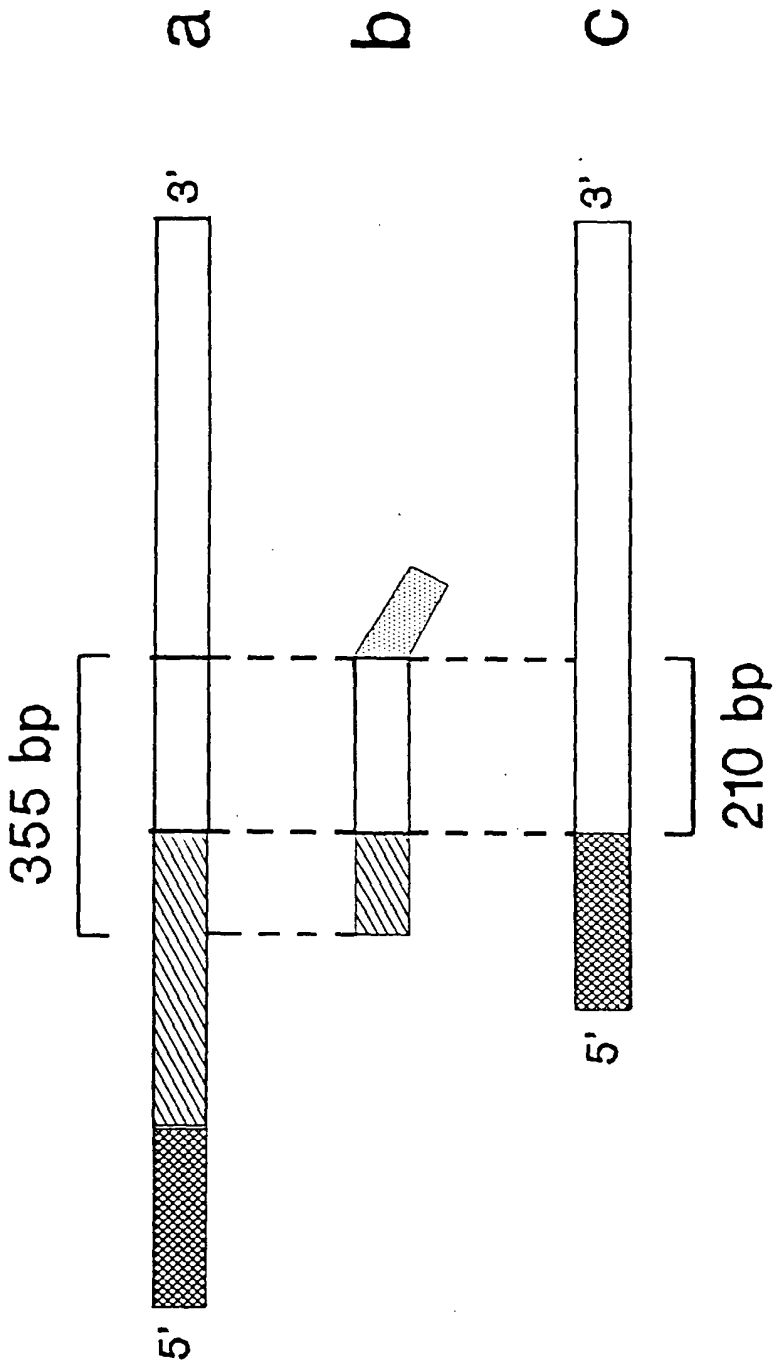


FIGURE 19

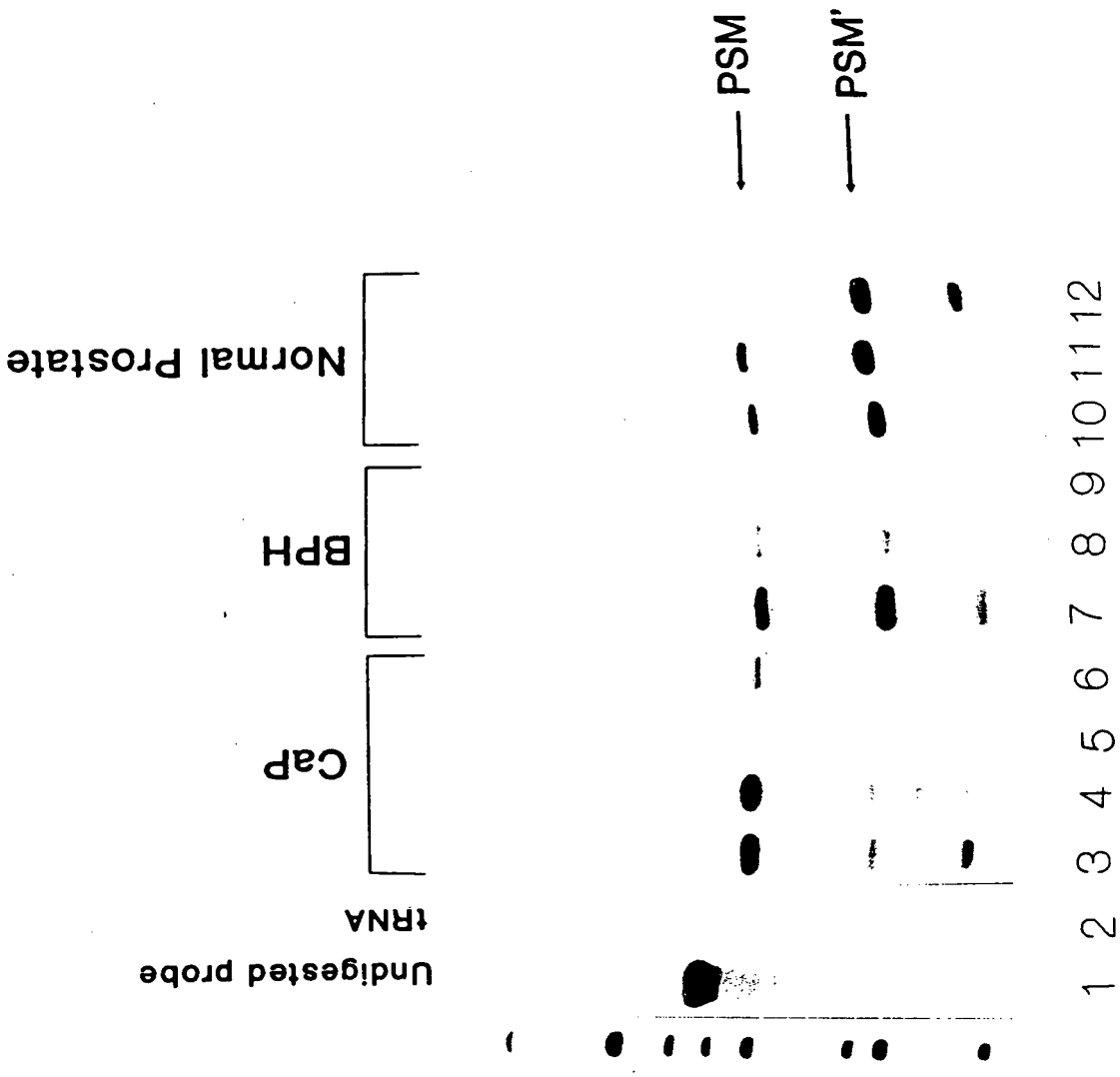

PSM
PSM'
Normal Prostate
BPH
CaP
tRNA
Undigested probe
1 2 3 4 5 6 7 8 9 10 11 12

FIGURE 20

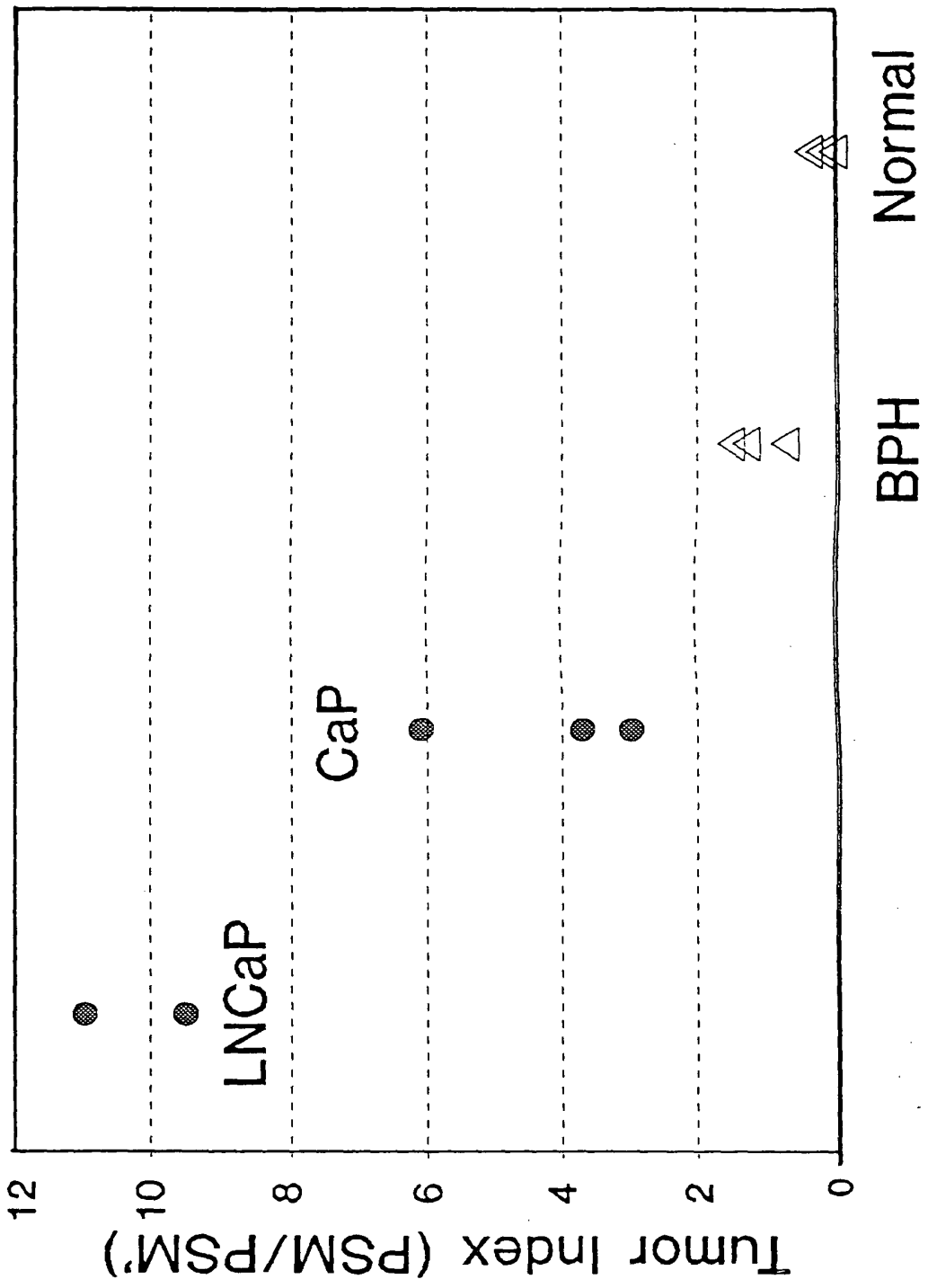


FIGURE 21

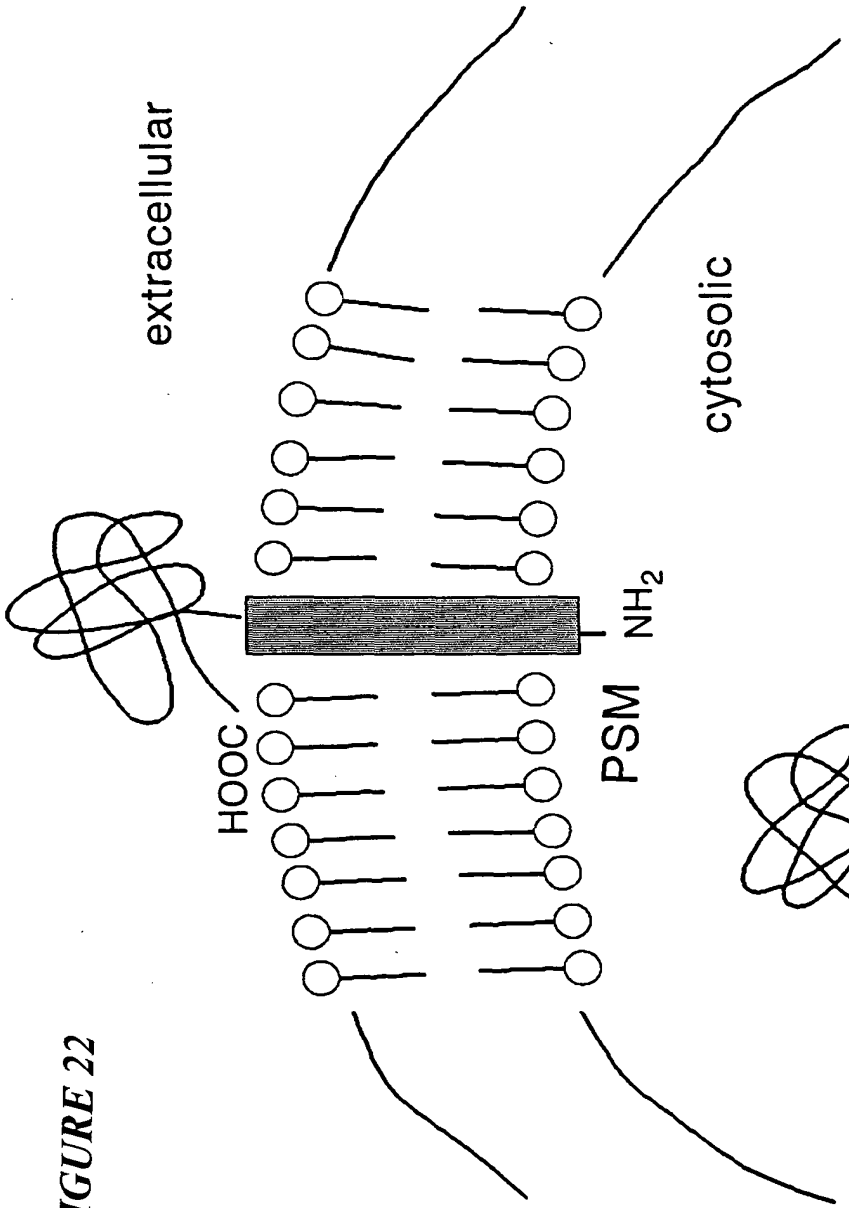
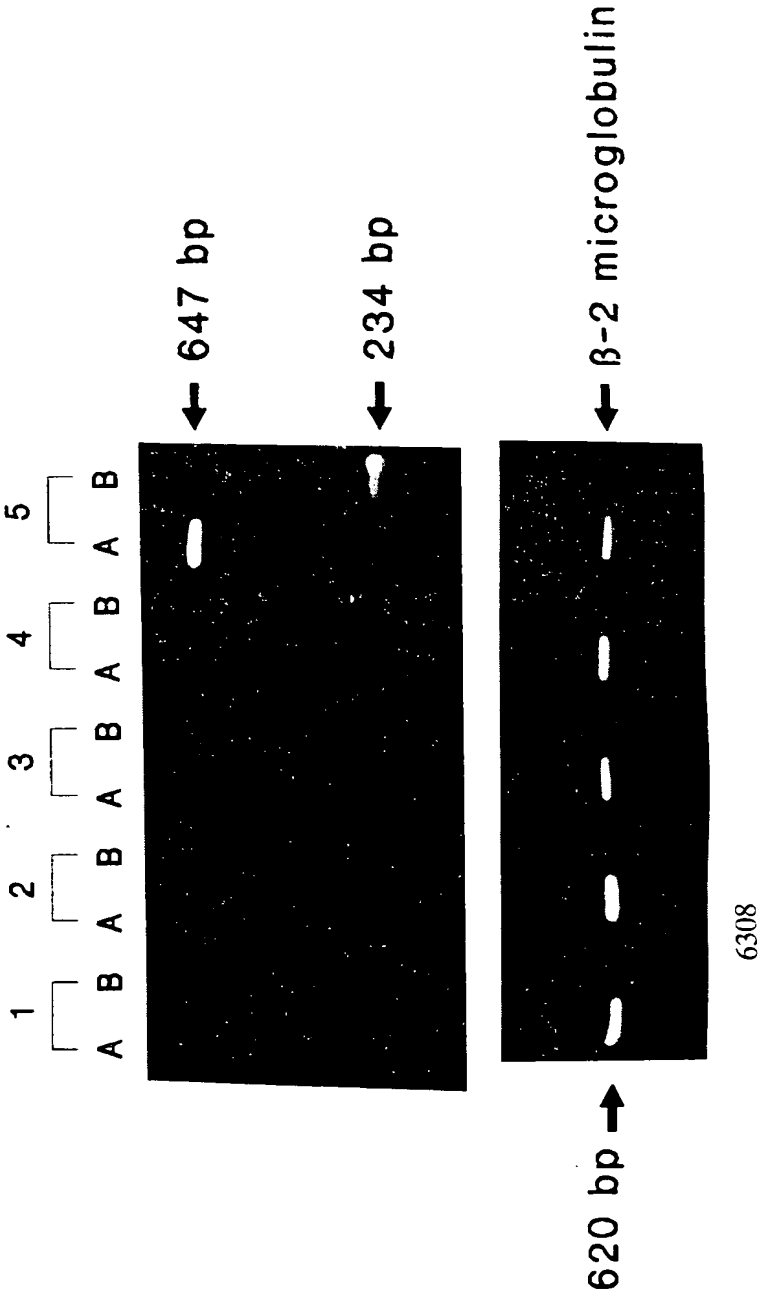


FIGURE 22

FIGURE 23



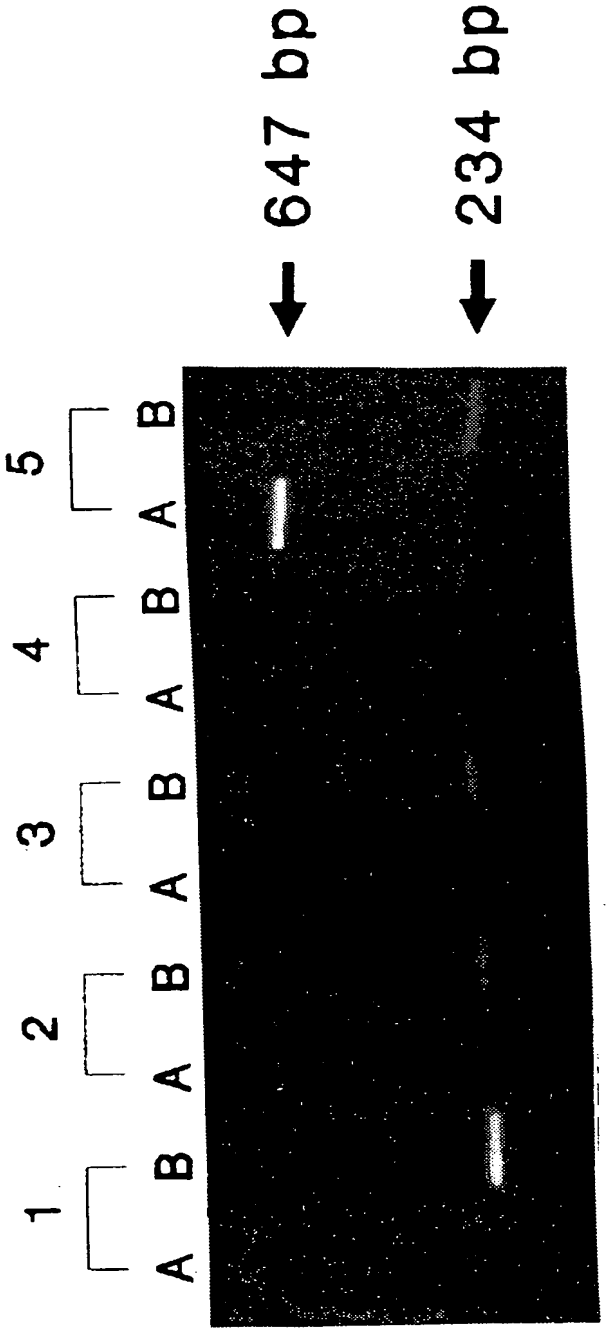


FIGURE 24

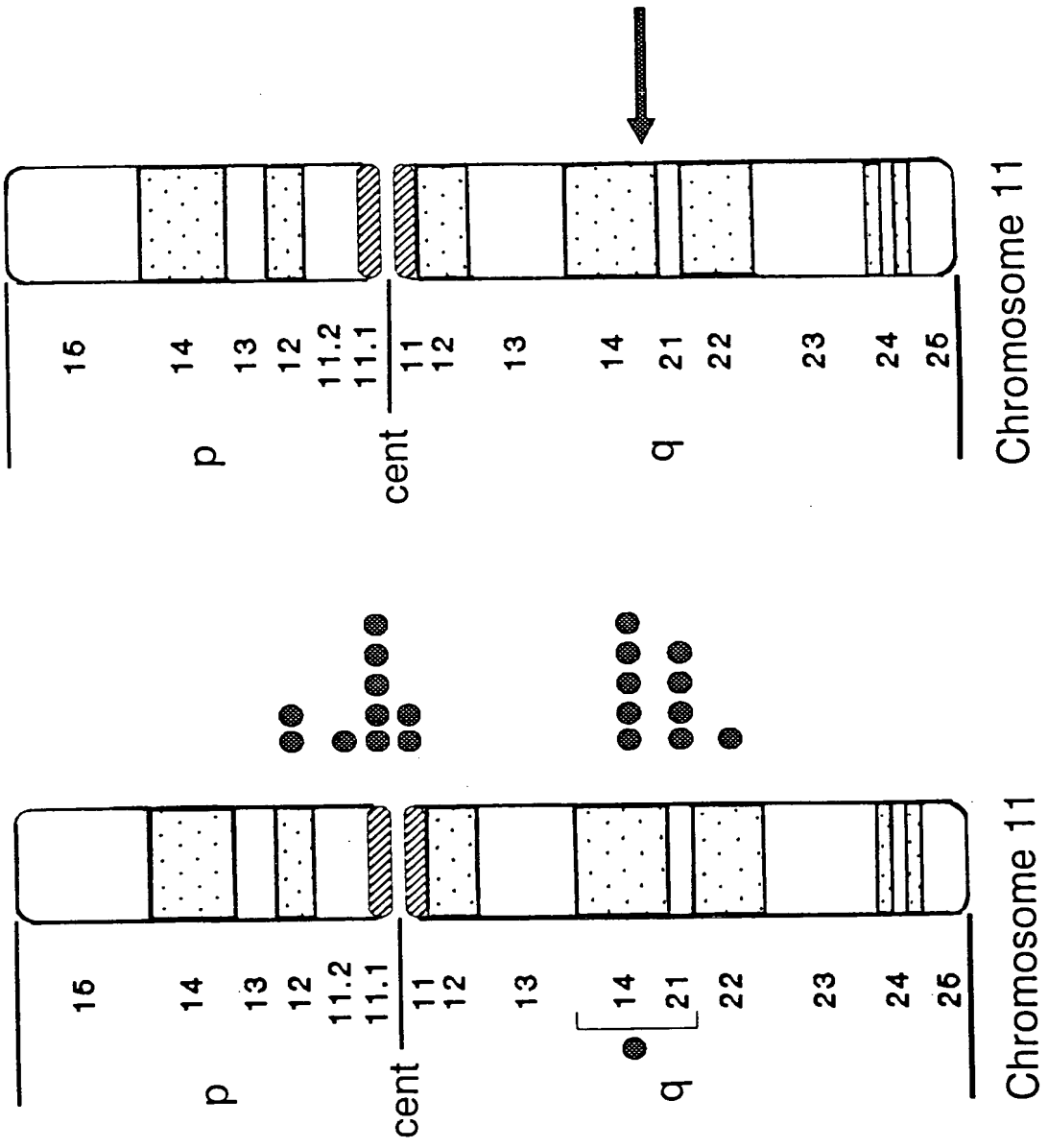

Cosmid
Chromosome 11
cDNA
Chromosome 11
p
cent
q
16
14
13
12
11.2
11.1
11
12
13
14
21
22
23
24
25

FIGURE 25

FIGURE 26

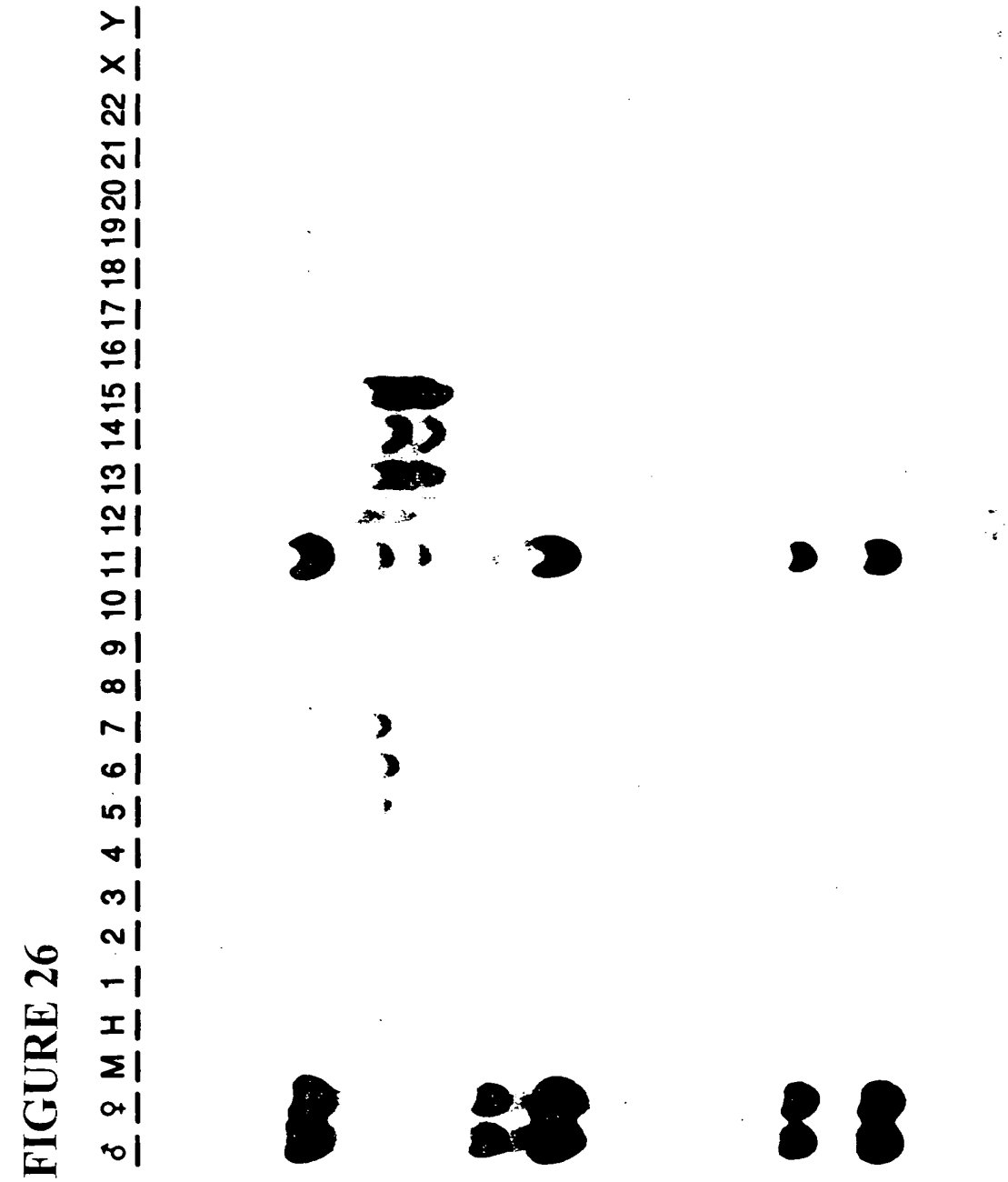
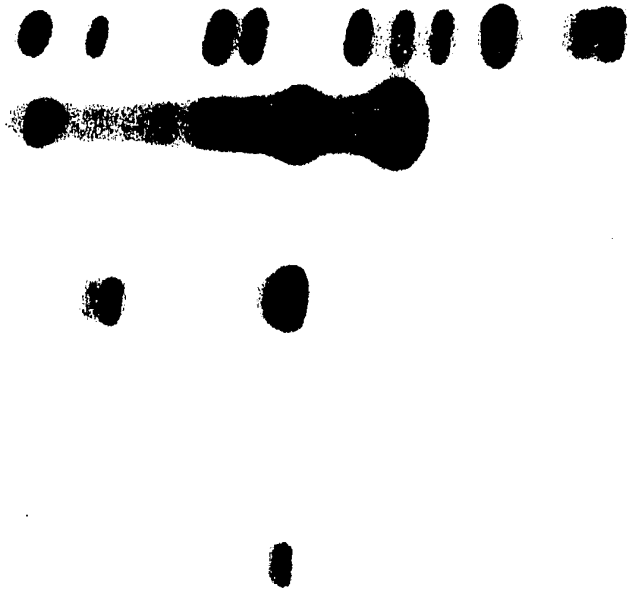


FIGURE 27



Markers

Uncut

t RNA

LnCap

PC 3

AT6.1

AT6.1-11 clone 1

AT6.1-11 clone 2

A9

A9 (11)

R1564

R1564-11 clone 4

R1564-11 clone 6

FIGURE 28

TISSUE/ CELL LINE	CANCER CELL TYPE	[1]PSM DNA	[2]PSM RNA
HUMAN PROSTATE	N.A.	+	+
HUMAN MAMMARY	N.A.	+	-
AT6.1	RAT PROSTATIC ADENOCARCINOMA	-	-
AT6.1-11-CL1	"	+	+
AT6.1-11-CL2	"	-	-
R1564	RAT MAMMARY ADENOCARCINOMA	-	-
R1564-11-CL2	"	+	-
R1564-11-CL4	"	+	-
R1564-11-CL5	"	+	-
R1564-11-CL6	"	+	-
A9	MOUSE FIBROSARCOMA	-	-
A9(11)	"	+	-

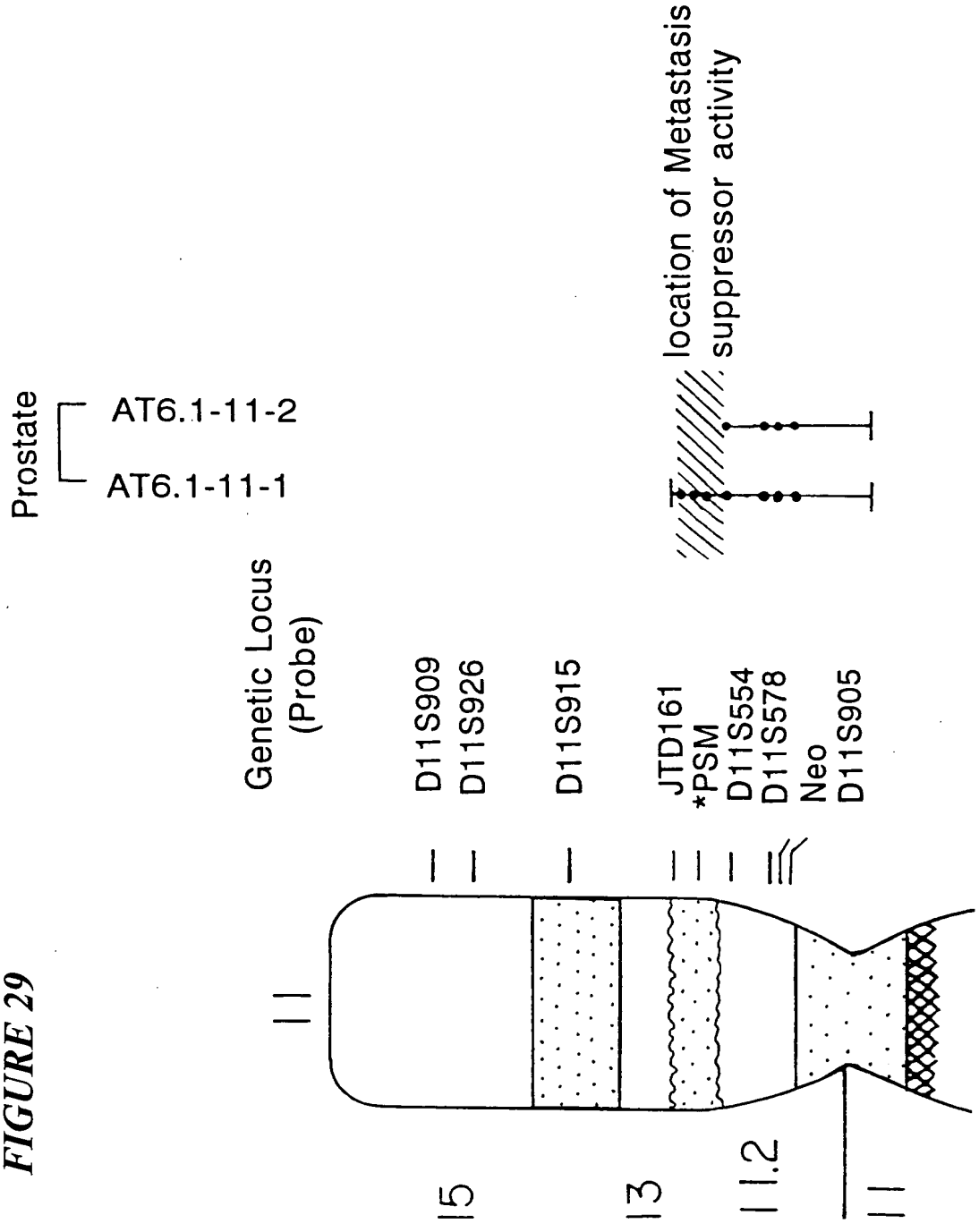

Prostate
AT6.1-11-2
AT6.1-11-1
location of Metastasis suppressor activity
Genetic Locus (Probe)
D11S909
D11S926
D11S915
JTD161
*PSM
D11S554
D11S578
Neo
D11S905
11
15
13
11.2
11

FIGURE 29

FIGURE 30

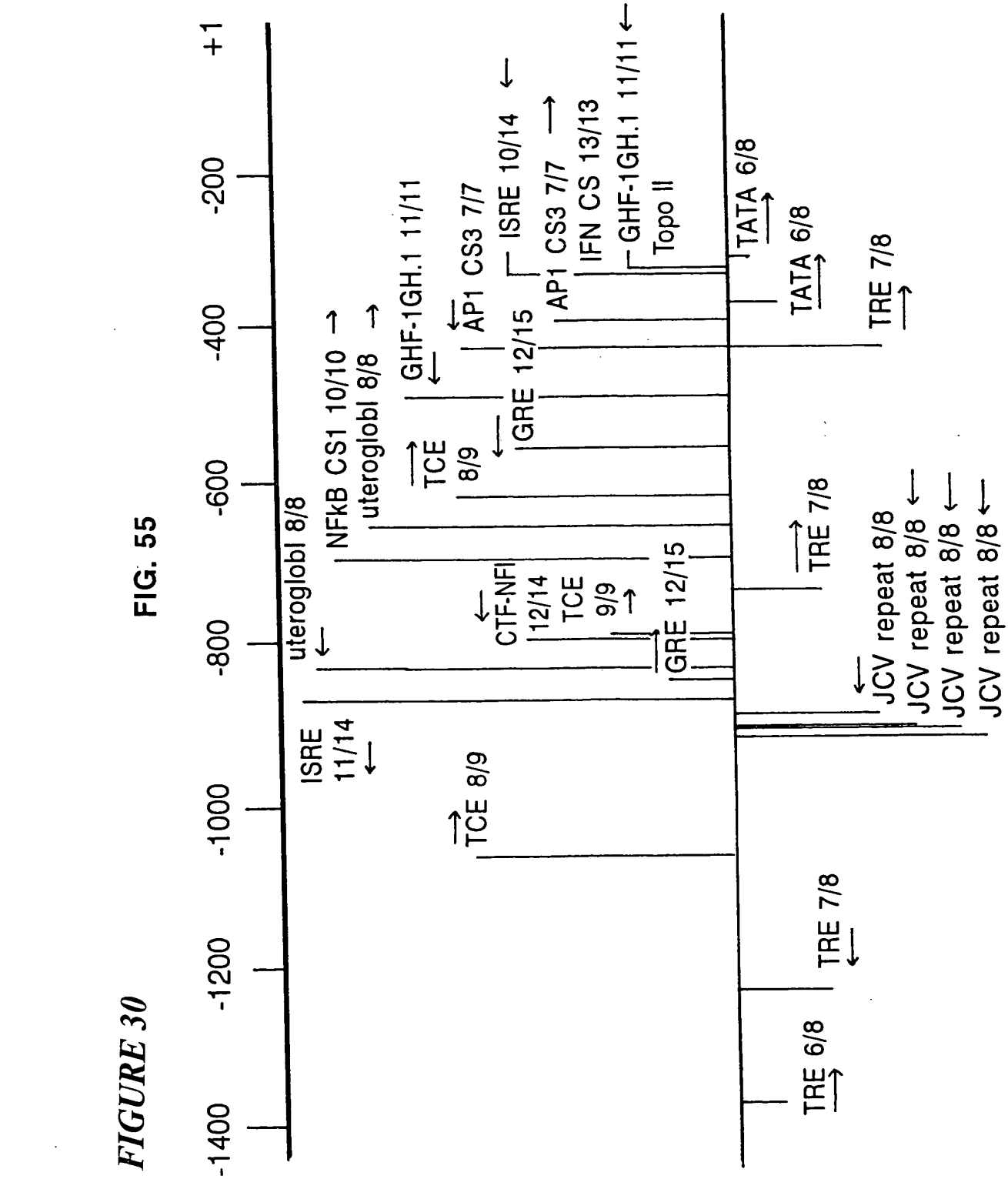


FIG. 55

FIGURE 31

Prostate Specific Promoter: Cytosine Deaminase Chimera

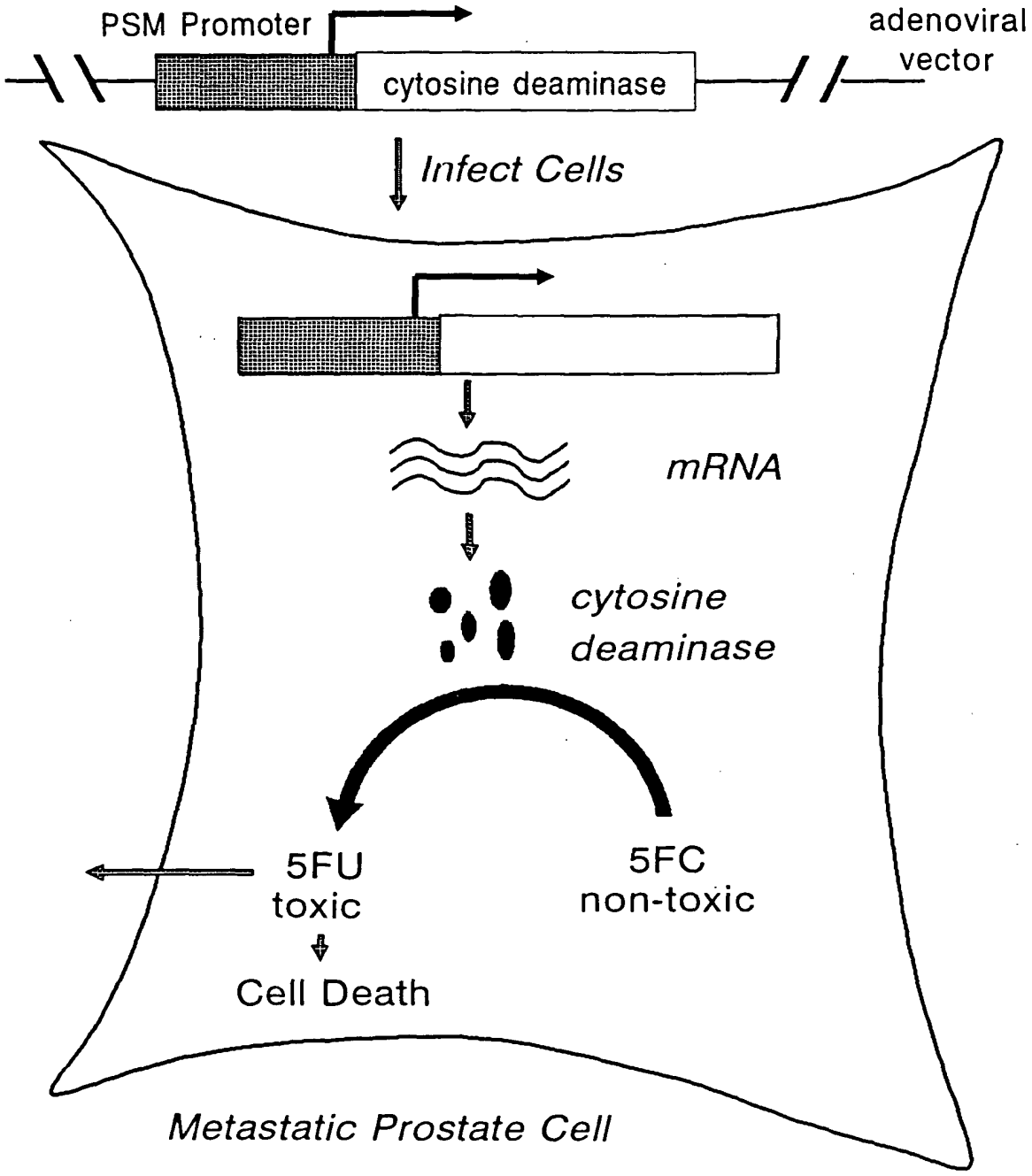


FIGURE 32A

```
          10         20         30         40         50         60
          |          |          |          |          |          |
  1 AAGGGTGCTC CTTAGGCTGA ATGCTTGCAG ACAGGATGCT TGGTTACAGA TGGGCTGTGA
    TTCCCACGAG GAATCCGACT TACGAACGTC TGTCCTACGA ACCAATGTCT ACCCGACACT

 61 CTCGAGTGGA GTTTTATAAG GGTGCTCCTT AGGCTGAATG CTTGCAGACA GGATGCTTGG
    GAGCTCACCT CAAAATATTC CCACGAGGAA TCCGACTTAC GAACGTCTGT CCTACGAACC

121 TTACAGATGG GCTGTGAGCT GGGTGCTTGT AAGAGGATGC TTGGGTGCTA AGTGAGCCAT
    AATGTCTACC CGACACTCGA CCCACGAACA TTCTCCTACG AACCCACGAT TCACTCGGTA

181 TTGCAGTTGA CCCTATTCTT GGAACATTCA TTCCCCTCTA CCCCTGTTTC TGTTCCTGCC
    AACGTCAACT GGGATAAGAA CCTTGTAAGT AAGGGGAGAT GGGGACAAAG ACAAGGACGG

241 AGCTAAGCCC ATTTTTCATT TTTCTTTTAA CTCCTTAGCG CTCCGCAAAA CTTAATCAAT
    TCGATTCGGG TAAAAAGTAA AAAGAAAATT GAGGAATCGC GAGGCGTTTT GAATTAGTTA

301 TTCTTTAAAC CTCAGTTTTC TTATCTGTAA AAGGTAAATA ATAATACAGG GTGCAACAGA
    AAGAAATTTG GAGTCAAAAG AATAGACATT TTCCATTTAT TATTATGTCC CACGTTGTCT

361 AAAATCTAGT GTGGTTTACA TAATCACCTG TTAGAGATTT TAAATTATTT CAGGATAAGT
    TTTTAGATCA CACCAAATGT ATTAGTGGAC AATCTCTAAA ATTTAATAAA GTCCTATTCA

421 CATGATAATT AAATGAAATA ATGCACATAA AGCACATAGT GTGGTGTCCT CCATATAGAA
    GTACTATTAA TTTACTTTAT TACGTGTATT TCGTGTATCA CACCACAGGA GGTATATCTT

481 AATGCTCAGT ATATTGGTTA TTAACTACTT GTTGAAGGTT TATCTTCTCC ACTAAACTGT
    TTACGAGTCA TATAACCAAT AATTGATGAA CAACTTCCAA ATAGAAGAGG TGATTTGACA

541 AAGTTCCACA AGCCTTACAA TATGTGACAG ATATTCATTC ATTGTCTGAA TTCTTCAAAT
    TTCAAGGTGT TCGGAATGTT ATACACTGTC TATAAGTAAG TAACAGACTT AAGAAGTTTA

601 ACATCCTCTT CACCATAGCG TCTTATTAAT TGAATTATTA ATTGAATAAA TTCTATTGTT
    TGTAGGAGAA GTGGTATCGC AGAATAATTA ACTTAATAAT TAACTTATTT AAGATAACAA

661 CAAAAATCAC TTTTATATTT AACTGAAATT TGCTTACTTA TAATCACATC TAACCTTCAA
    GTTTTTAGTG AAAATATAAA TTGACTTTAA ACGAATGAAT ATTAGTGTAG ATTGGAAGTT

721 AGAAAACACA TTAACCAACT GTACTGGGTA ATGTTACTGG GTGATCCCAC GTTTTACAAA
    TCTTTTGTGT AATTGGTTGA CATGACCCAT TACAATGACC CACTAGGGTG CAAAATGTTT
```

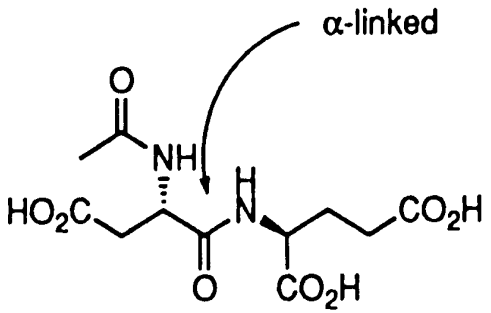
FIGURE 32B

```
 781 TGAGAAGATA TATTCTGGTA AGTTGAATAC TTAGCACCCA GGGGTAATCA GCTTGGACAG
     ACTCTTCTAT ATAAGACCAT TCAACTTATG AATCGTGGGT CCCCATTAGT CGAACCTGTC

 841 GACCAGGTCC AAAGACTGTT AAGAGTCTTC TGACTCCAAA CTCAGTGCTC CCTCCAGTGC
     CTGGTCCAGG TTTCTGACAA TTCTCAGAAG ACTGAGGTTT GAGTCACGAG GGAGGTCACG

 901 CACAAGCAAA CTCCATAAAG GTATCCTGTG CTGAATAGAG ACTGTAGAGT GGTACAAAGT
     GTGTTCGTTT GAGGTATTTC CATAGGACAC GACTTATCTC TGACATCTCA CCATGTTTCA

 961 AAGACAGACA TTATATTAAG TCTTAGCTTT GTGACTTCGA ATGACTTACC TAATCTAGCT
     TTCTGTCTGT AATATAATTC AGAATCGAAA CACTGAAGCT TACTGAATGG ATTAGATCGA

1021 AAATTTCAGT TTTACCATGT GTAAATCAGG AAGAGTAATA GAACAAACCT TGAAGGGTCC
     TTTAAAGTCA AAATGGTACA CATTTAGTCC TTCTCATTAT CTTGTTTGGA ACTTCCCAGG

1081 CAATGGTGAT TAAATGAGGT GATGTACATA ACATGCATCA CTCATAATAA GTGCTCTTTA
     GTTACCACTA ATTTACTCCA CTACATGTAT TGTACGTAGT GAGTATTATT CACGAGAAAT

1141 AATATTAGTC ACTATTATTA GCCATCTCTG ATTAGATTTG ACAATAGGAA CATTAGGAAA
     TTATAATCAG TGATAATAAT CGGTAGAGAC TAATCTAAAC TGTTATCCTT GTAATCCTTT

1201 GATATAGTAC ATTCAGGATT TTGTTAGAAA GAGATGAAGA AATTCCCTTC CTTCCTGCCC
     CTATATCATG TAAGTCCTAA AACAATCTTT CTCTACTTCT TTAAGGGAAG GAAGGACGGG

1261 TAGGTCATCT AGGAGTTGTC ATGGTTCATT GTTGACAAAT TAATTTTCCC AAATTTTTCA
     ATCCAGTAGA TCCTCAACAG TACCAAGTAA CAACTGTTTA ATTAAAAGGG TTTAAAAAGT

1321 CTTTGCTCAG AAAGTCTACA TCGAAGCACC CAAGACTGTA CAATCTAGTC CATCTTTTTC
     GAAACGAGTC TTTCAGATGT AGCTTCGTGG GTTCTGACAT GTTAGATCAG GTAGAAAAAG

1381 CACTTAACTC ATACTGTGCT CTCCCTTTCT CAAAGCAAAC TGTTTGCTAT TCCTTGAATA
     GTGAATTGAG TATGACACGA GAGGGAAAGA GTTTCGTTTG ACAAACGATA AGGAACTTAT

1441 CACTCTGAGT TTTCTGCCTT TGCCTACTCA GCTGGCCCAT GGCCCCTAAT GTTTCTTCTC
     GTGAGACTCA AAAGACGGAA ACGGATGAGT CGACCGGGTA CCGGGGATTA CAAAGAAGAG

1501 ATCTCCACTG GGTCAAATCC TACCTGTACC TTATGGTTCT GTTAAAAGCA GTGCTTCCAT
     TAGAGGTGAC CCAGTTTAGG ATGGACATGG AATACCAAGA CAATTTTCGT CACGAAGGTA

1561 AAAGTACTCC TAGCAAATGC ACGGCCTCTC TCACGGATTA TAAGAACACA GTTTATTTTA
```

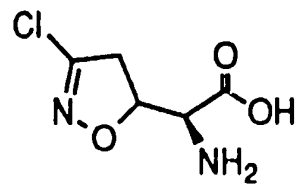
FIGURE 32C

```
    TTTCATGAGG ATCGTTTACG TGCCGGAGAG AGTGCCTAAT ATTCTTGTGT CAAATAAAAT

1621 TAAAGCATGT AGCTATTCTC TCCCTCGAAA TACGATTATT ATTATTAAGA ATTTATAGCA
     ATTTCGTACA TCGATAAGAG AGGGAGCTTT ATGCTAATAA TAATAATTCT TAAATATCGT

1681 GGGATATAAT TTTGTATGAT GATTCTTCTG GTTAATCCAA CCAAGATTGA TTTTATATCT
     CCCTATATTA AAACATACTA CTAAGAAGAC CAATTAGGTT GGTTCTAACT AAAATATAGA

1741 ATTACGTAAG ACAGTAGCCA GACATAGCCG GGATATGAAA ATAAAGTCTC TGCCTTCAAC
     TAATGCATTC TGTCATCGGT CTGTATCGGC CCTATACTTT TATTTCAGAG ACGGAAGTTG

1801 AAGTTCCAGT ATTCTTTTCT TTCCTCCCCT CCCCTCCCCT CCCTTCCCCT CCCCTTCCTT
     TTCAAGGTCA TAAGAAAAGA AAGGAGGGGA GGGGAGGGGA GGGAAGGGGA GGGGAAGGAA

1861 CCCTTTCCCT TCCCTTCCTT TCTTTCTTGA GGGAGTCTCA CTCTGTCACC AGGCTCCAGT
     GGGAAAGGGA AGGGAAGGAA AGAAAGAACT CCCTCAGAGT GAGACAGTGG TCCGAGGTCA

1921 GCAGTGGCGC TATCTTGGCT GACTGCAACC TCCGCCTCCC CGGTTCAAGC GATTCTCCTG
     CGTCACCGCG ATAGAACCGA CTGACGTTGG AGGCGGAGGG GCCAAGTTCG CTAAGAGGAC

1981 CCTCAGCCTC CTGAGTAGCT GGGACTACAG GAGCCCGCCA CCACGCCCAG CTAATTTTTG
     GGAGTCGGAG GACTCATCGA CCCTGATGTC CTCGGGCGGT GGTGCGGGTC GATTAAAAAC

2041 TATTTTTAGT AGAGATGGGG TTTCACCATG TTGGCCAGGA TGGTCTCGAT TTCTCGACTT
     ATAAAAATCA TCTCTACCCC AAAGTGGTAC AACCGGTCCT ACCAGAGCTA AAGAGCTGAA

2101 CGTGATCCGC CTGTCTGGGC CTCCCAAAGT GCTGGGATTA CAGGCGTGAG CCACCACGCC
     GCACTAGGCG GACAGACCCG GAGGGTTTCA CGACCCTAAT GTCCGCACTC GGTGGTGCGG

2161 CGGCTTTAAA AAATGGTTTT GTAATGTAAG TGGAGGATAA TACCCTACAT GTTTATTAAT
     GCCGAAATTT TTTACCAAAA CATTACATTC ACCTCCTATT ATGGGATGTA CAAATAATTA

2221 AACAATAATA TTCTTTAGGA AAAAGGGCGC GGTGGTGATT TACACTGATG ACAAGCATTC
     TTGTTATTAT AAGAAATCCT TTTTCCCGCG CCACCACTAA ATGTGACTAC TGTTCGTAAG

2281 CCGACTATGG AAAAAAAGCG CAGCTTTTTC TGCTCTGCTT TTATTCAGTA GAGTATTGTA
     GGCTGATACC TTTTTTTCGC GTCGAAAAAG ACGAGACGAA AATAAGTCAT CTCATAACAT

2341 GAGATTGTAT AGAATTTCAG AGTTGAATAA AAGTTCCTCA TAATTATAGG AGTGGAGAGA
     CTCTAACATA TCTTAAAGTC TCAACTTATT TTCAAGGAGT ATTAATATCC TCACCTCTCT
```

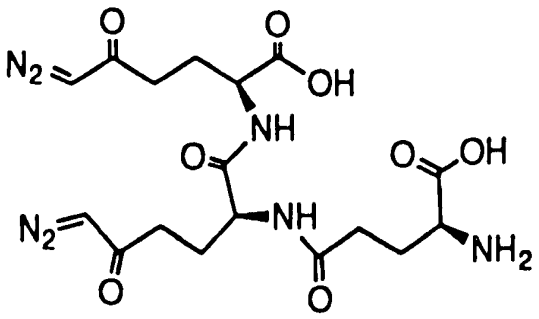
FIGURE 33



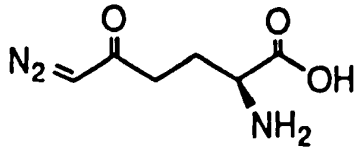
NAAG 1
N-acetylaspartyl-L-glutamate



Acivicin

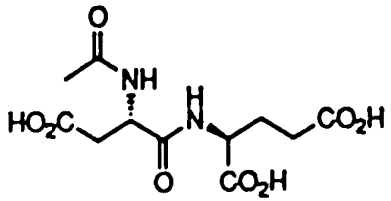


Azotomycin, becomes active by *in vivo* conversion to DON

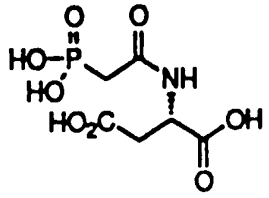


6-diazo-5-oxo-norleucine, DON

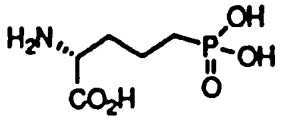
FIGURE 34



NAAG 1
N-acetylaspartyl-L-glutamate

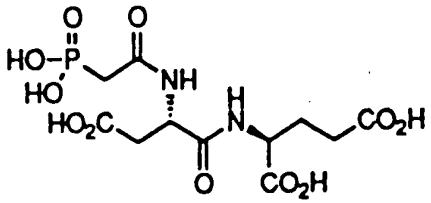


PALA 2
N-phosphonoacetyl-L-aspartate

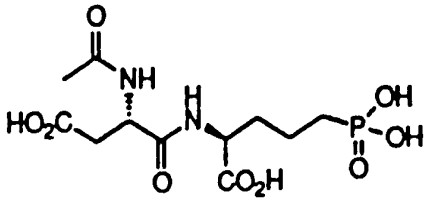


D-APV

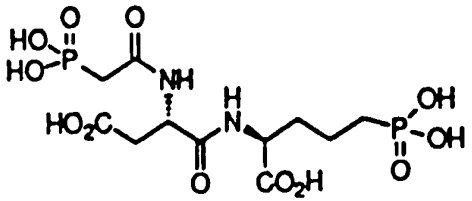
L-Glu antagonist



3

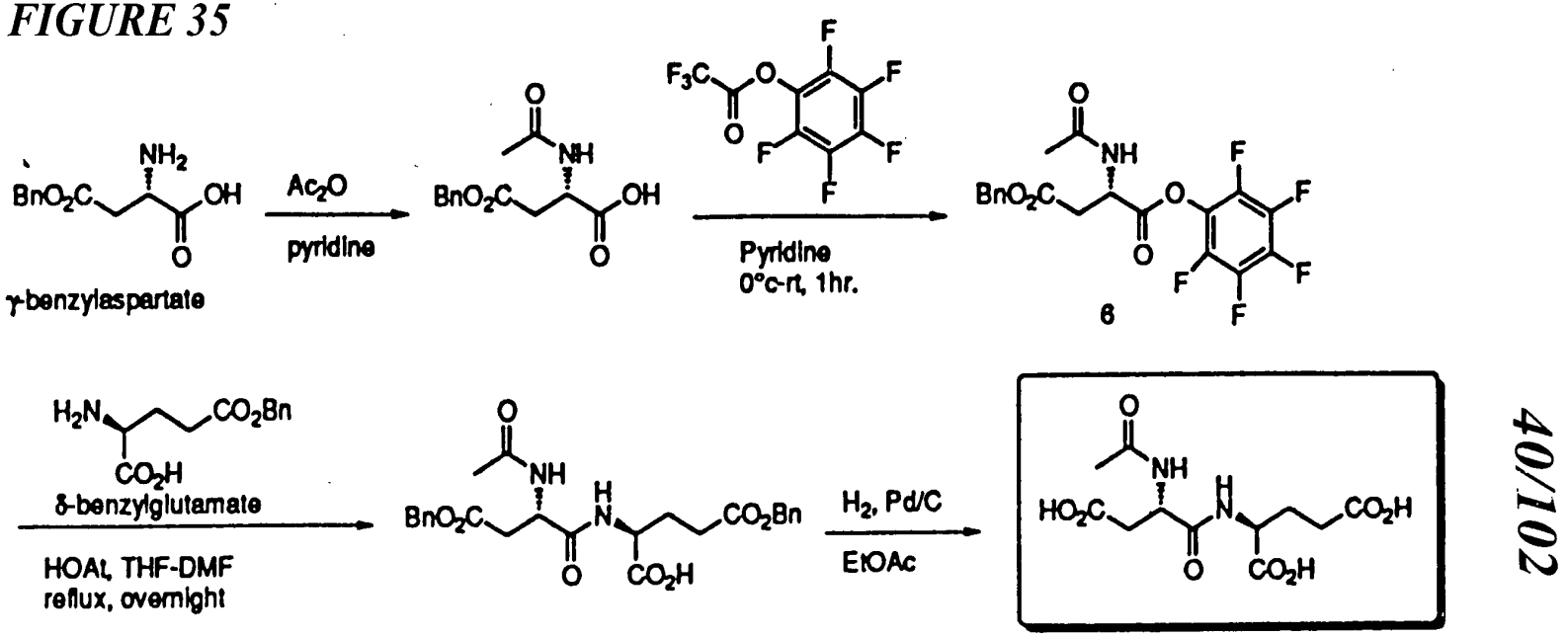


4



5

FIGURE 35



Ac_2O = acetic anhydride
THF = t trahydrofurane
DMF = N,N-dimethylformamide
Pd/C = palladium on charcoal
EtOAc = ethylacetate

HOAt =

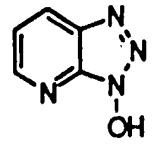


FIGURE 36

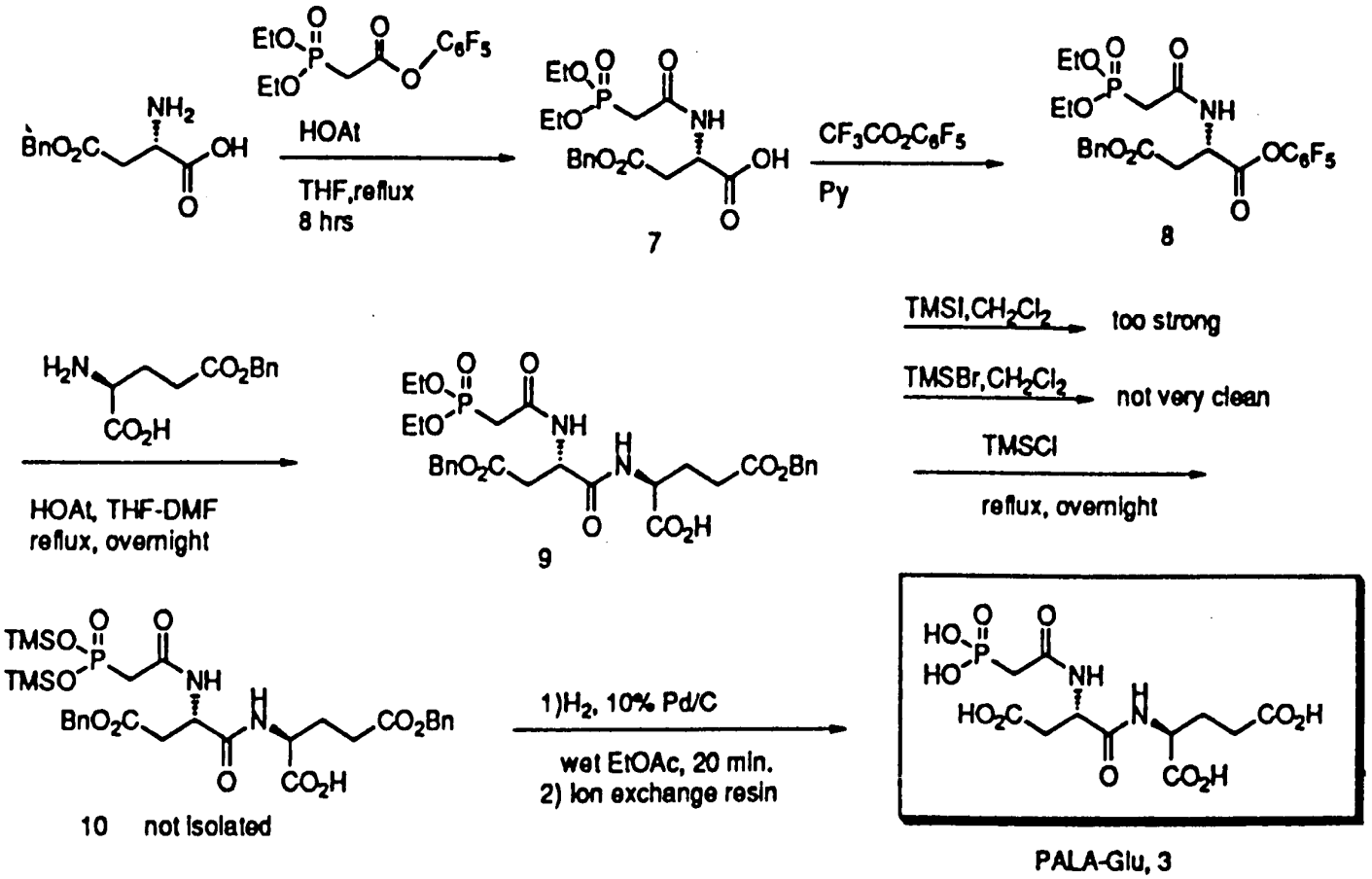

HOAt
THF,reflux
8 hrs
7
$CF_3CO_2C_6F_5$
Py
8
TMSI,CH_2Cl_2
too strong
TMSBr,CH_2Cl_2
not very clean
TMSCl
reflux, overnight
HOAt, THF-DMF
reflux, overnight
9
10 not isolated
1)H_2, 10% Pd/C
wet EtOAc, 20 min.
2) Ion exchange resin
PALA-Glu, 3

FIGURE 37

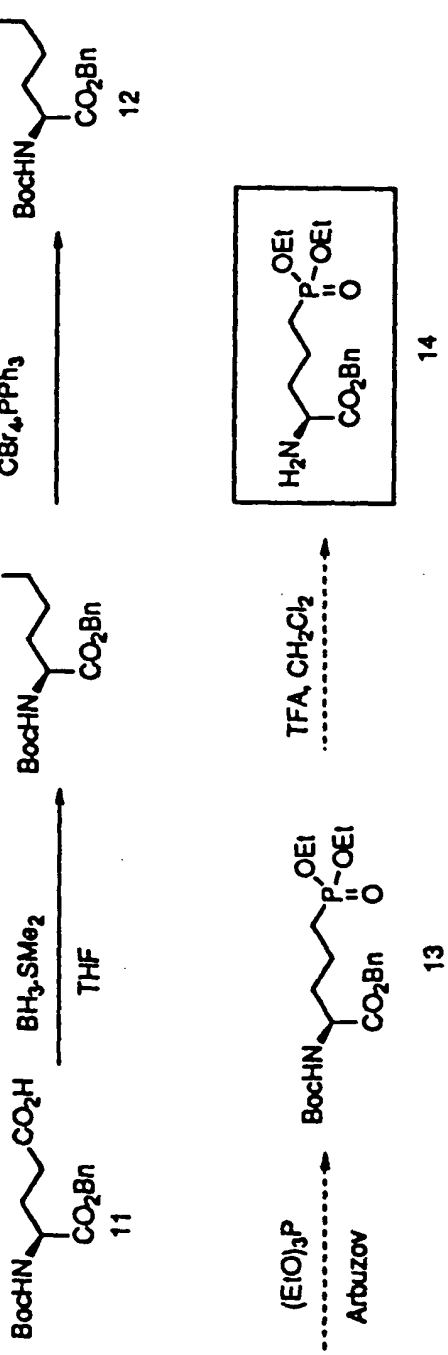

BocHN
CO2Bn
CO2H
11
BH3.SMe2
THF
BocHN
CO2Bn
OH
CBr4,PPh3
BocHN
CO2Bn
Br
12
(EtO)3P
Arbuzov
BocHN
CO2Bn
P
OEt
OEt
O
13
TFA, CH2Cl2
H2N
CO2Bn
P
OEt
OEt
O
14

FIGURE 38

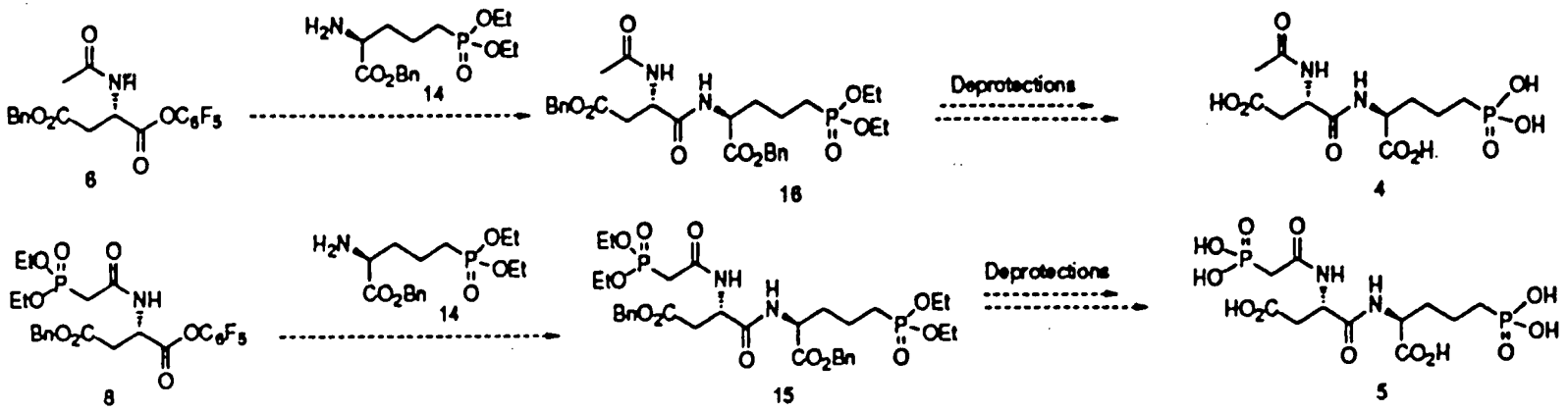
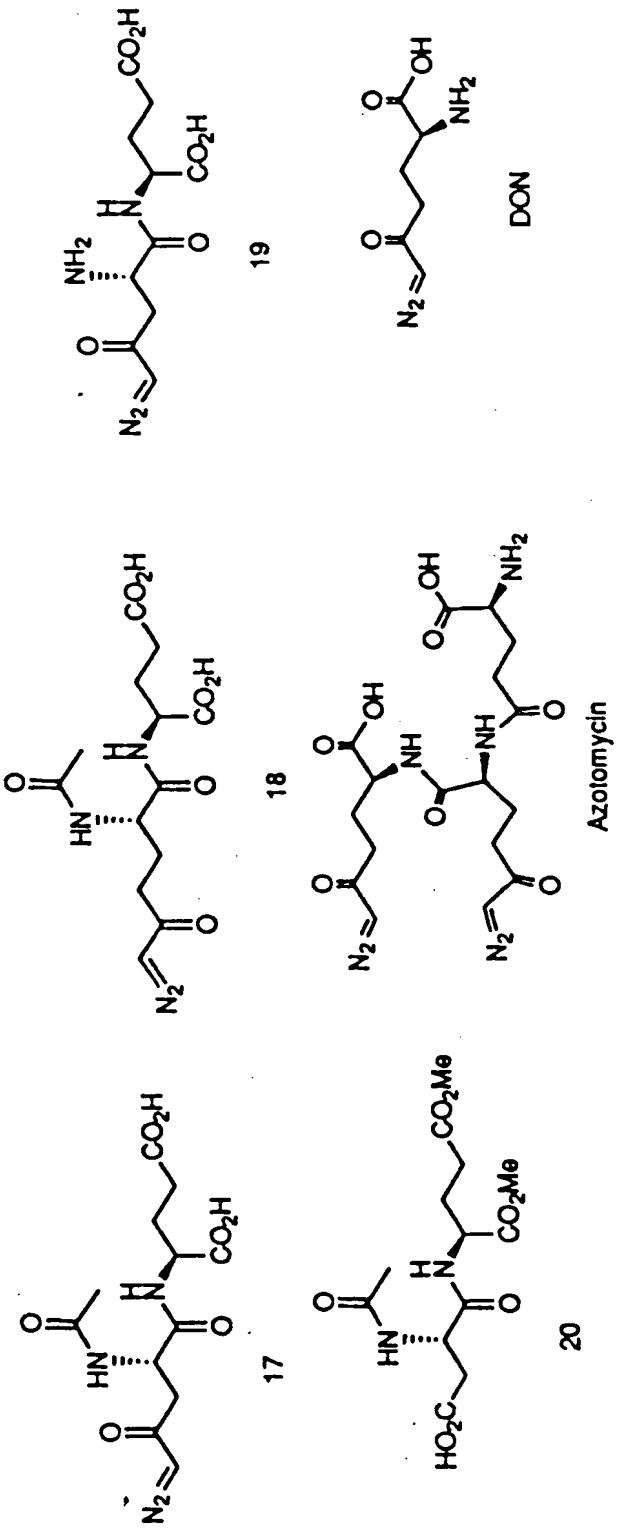
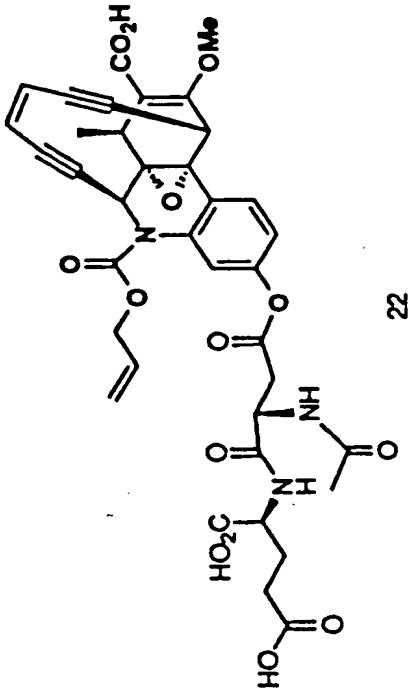
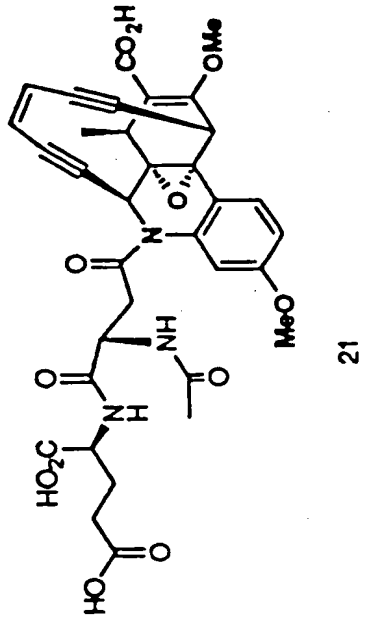


FIGURE 39





22



21

FIGURE 40

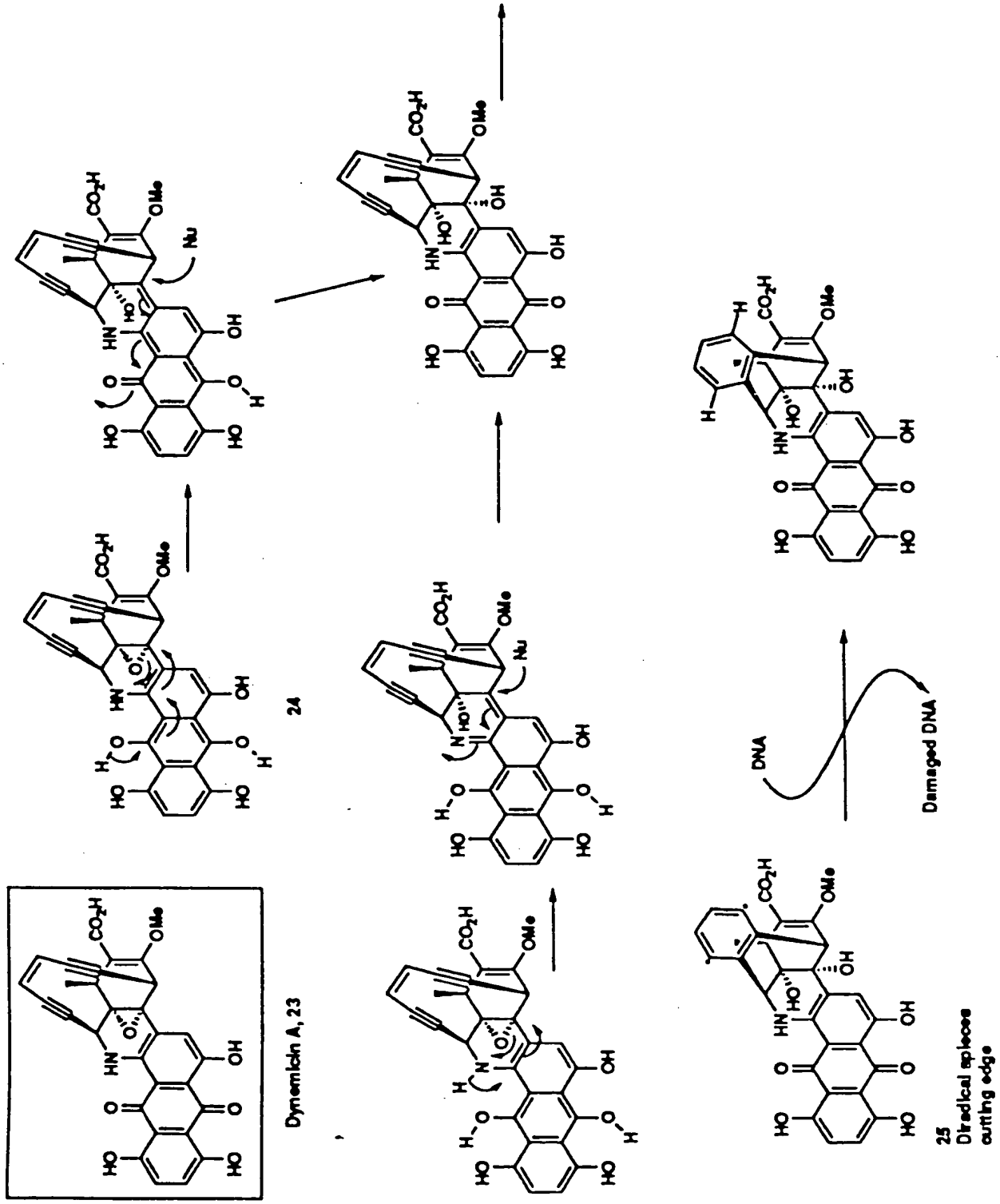
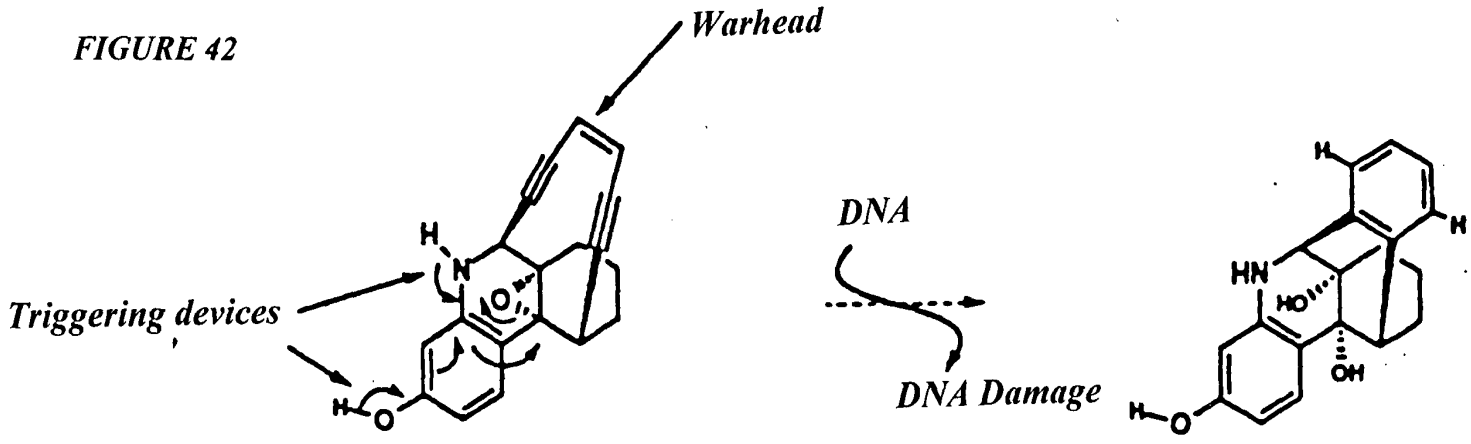

Dynemicin A, 23
24
25
Diradical species
cutting edge
DNA
Damaged DNA

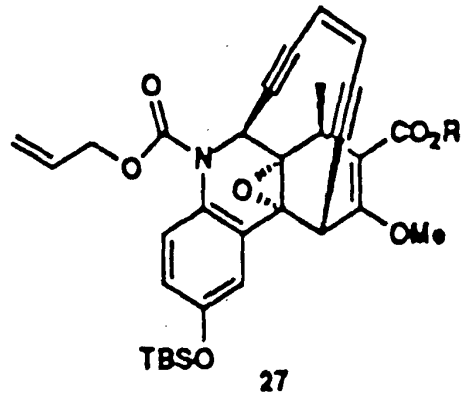
FIGURE 41

FIGURE 42

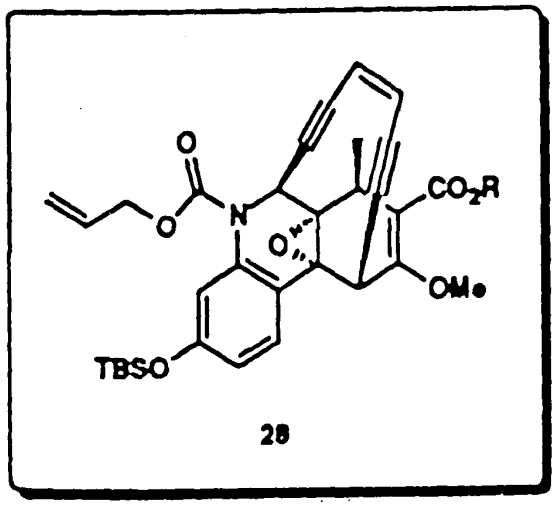


26

active at the nano to picomolar levels in different cell lines readily rearranges when one or both triggering devices are deprotected



27



28

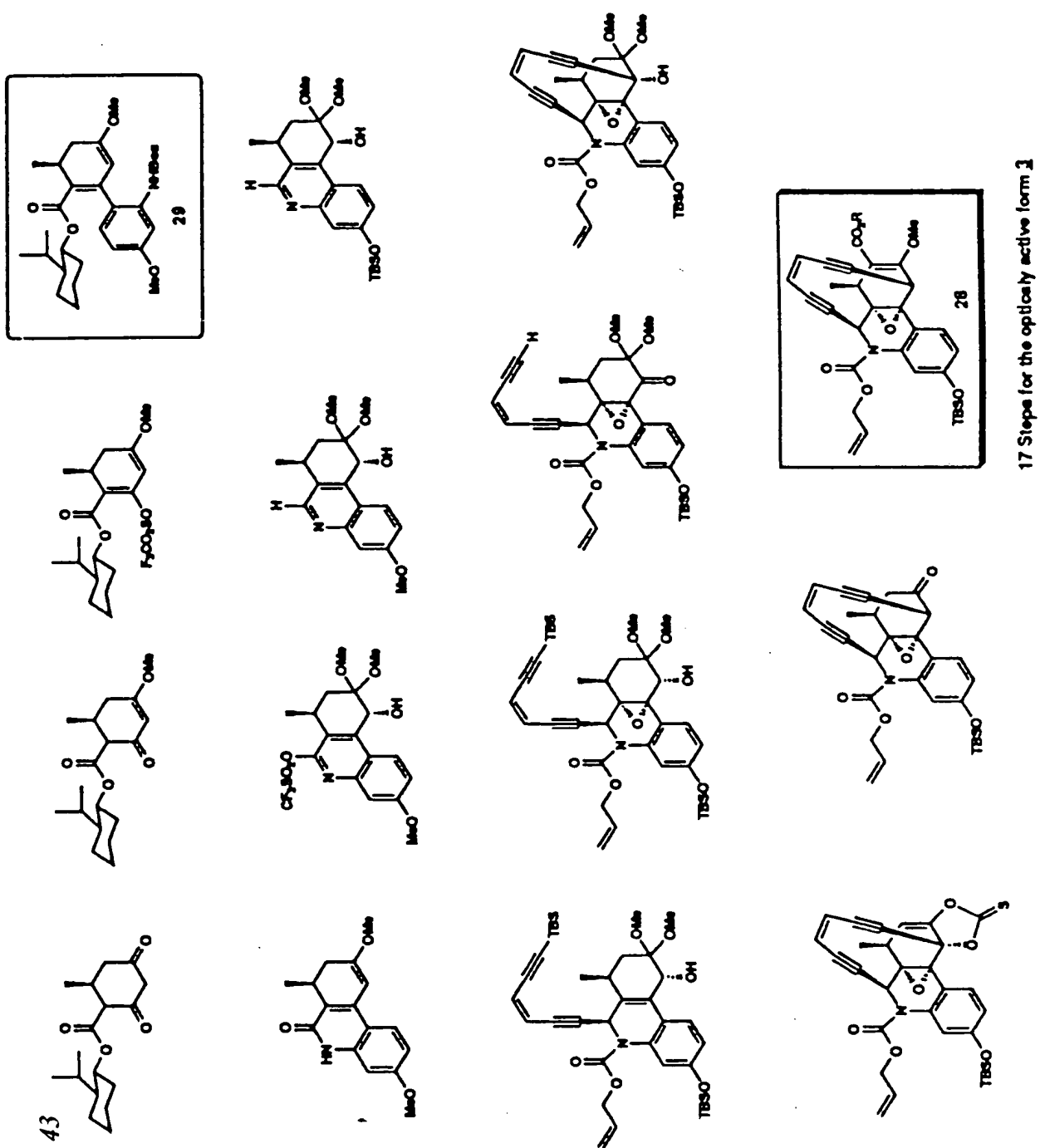


FIGURE 43

FIGURE 44

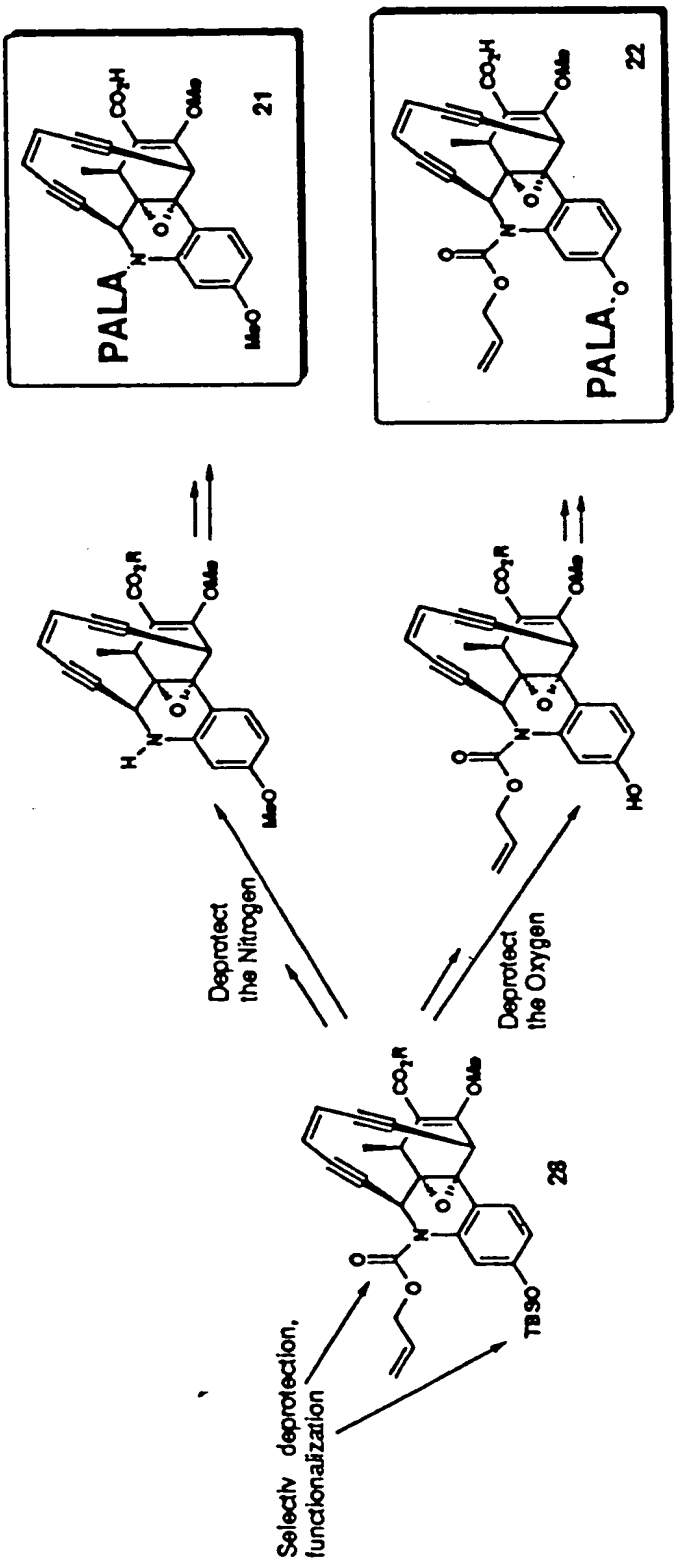

PALA
21
PALA
22
Deprotect the Nitrogen
Deprotect the Oxygen
Selectiv deprotection, functionalization
28

FIGURE 45

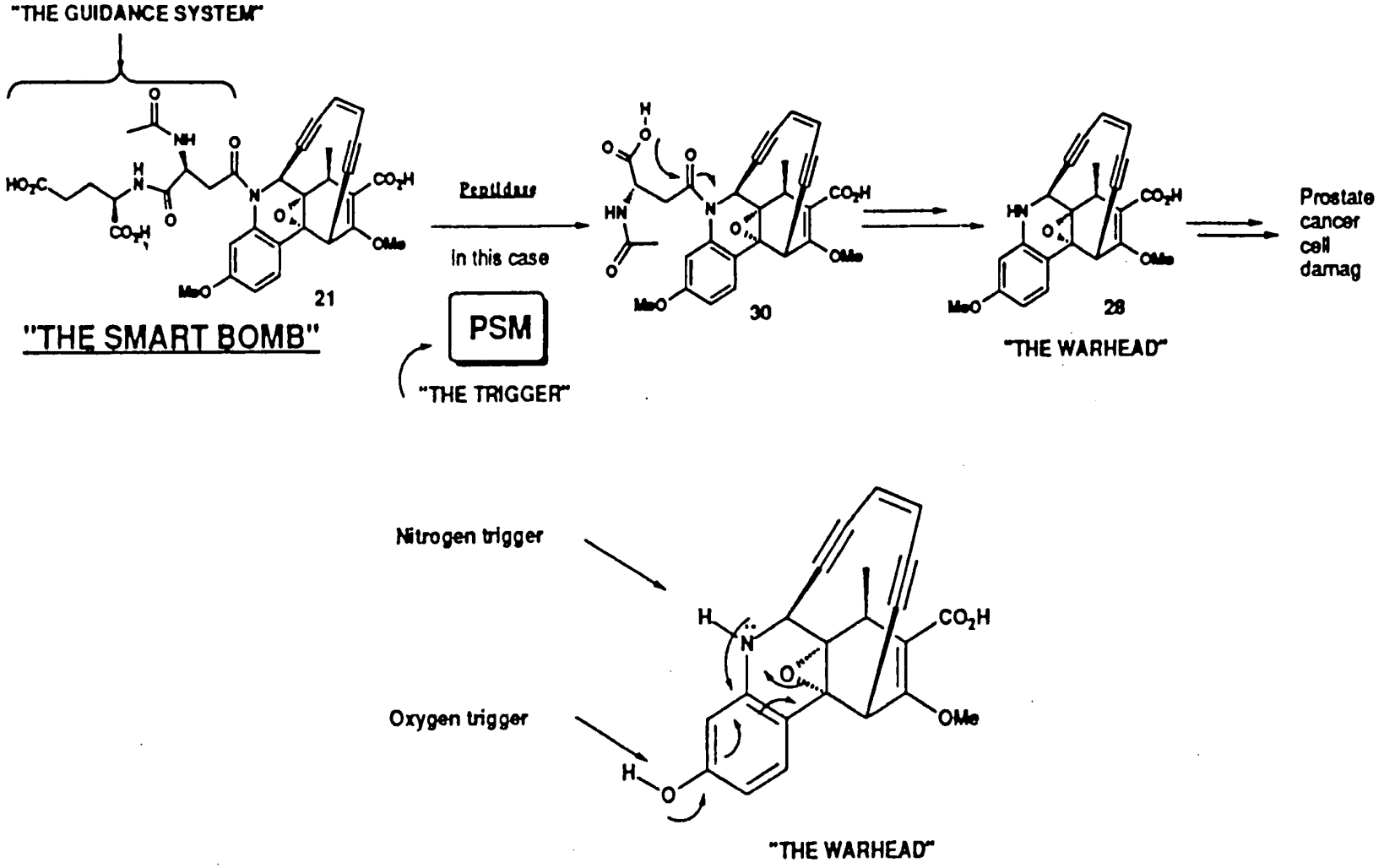

"THE GUIDANCE SYSTEM"
21
"THE SMART BOMB"
Peptidase
In this case
PSM
"THE TRIGGER"
30
28
"THE WARHEAD"
Prostate cancer cell damag
Nitrogen trigger
Oxygen trigger
"THE WARHEAD"

FIGURE 46A

```
                10         20         30         40         50         60
                |          |          |          |          |          |
  1 TAGGGGGGCG CCTCGCGGAG AAACCTCGGA GTCTTCCCCG TGGTGCCGCG GTGCTGGGAC
    ATCCCCCCGC GGAGCGCCTC TTTGGAGCCT CAGAAGGGGC ACCACGGCGC CACGACCCTG

 61 TCGCGGGTCA GCTGCCGAGT GGGATCCTGT TGCTGGTCTT CCCCAGGGGC GGCGATTAGG
    AGCGCCCAGT CGACGGCTCA CCCTAGGACA ACGACCAGAA GGGGTCCCCG CCGCTAATCC

121 GTCGGGGTAA TGTGGGGTGA GCACCCCTCG AGTTAGGAGG AGGGTAGCTG GGAACGGTGC
    CAGCCCCATT ACACCCCACT CGTGGGGAGC TCAATCCTCC TCCCATCGAC CCTTGCCACG

181 AGGGCTGAGT TCTCGACAAG CTGCTGGTAG GACAGTCACT CAGGTTGAGG GTAGAACTGA
    TCCCGACTCA AGAGCTGTTC GACGACCATC CTGTCAGTGA GTCCAACTCC CATCTTGACT

241 GAGAACCTGA AACTGGGCGT AGGAAGGTTC CAAGTGCTGG AGCCCTGCAA GACAGAGGAA
    CTCTTGGACT TTGACCCGCA TCCTTCCAAG GTTCACGACC TCGGGACGTT CTGTCTCCTT

301 GTTTTTTTTT TGCTTTTGTT TTGTTTTGTT TTGTTTTGTT TTGTTTTGTT TGTTTGTTTG
    CAAAAAAAAA ACGAAAACAA AACAAAACAA AACAAAACAA AACAAAACAA ACAAACAAAC

361 TTTTTTTACC TCTCTGTGCA TTCTTTCTTC CTTGGAAGTA ACAGAGGCAA GCTTGGGAAC
    AAAAAAATGG AGAGACACGT AAGAAAGAAG GAACCTTCAT TGTCTCCGTT CGAACCCTTG

421 TGTGTGAACC AGGTCAGCAA TCTGCACAGG TCTTTACCAG CGGGTCTTTT GCTGTTTTTC
    ACACACTTGG TCCAGTCGTT AGACCTGTCC AGAAATGGTC GCCCAGAAAA CGACAAAAAG

481 CTGGGTACTG ATTTGCAGAC TTGATCCAAC TTTCTAAGAA AAGCAGAACC ACACAGGCAA
    GACCCATGAC TAAACGTCTG AACTAGGTTG AAAGATTCTT TTCGTCTTGG TGTGTCCGTT

541 GCTCAGACTC TTTTATTAAA TTCCAGTTTT GACTTTGCCA CTTCTTAGTG GCCTTGAACA
    CGAGTCTGAG AAAATAATTT AAGGTCAAAA CTGAAACGGT GAAGAATCAC CGGAACTTGT
```

FIGURE 46B

```
501 AGTTACCGAC TCCCTCTCAG CGTTAGTTAC CCTATTTTAT GATGAGGATA ATATTATCTG
    TCAATGGCTC AGGGAGAGTC GCAATCAATG GGATAAAATA CTACTCCTAT TATAATAGAC

561 CAAATTATTG GTAATAGTAA ATAATATAGC ATGTAAATCT CCTAGCACAG TACTGGGATT
    GTTTAATAAC CATTATCATT TATTATATCG TACATTTAGA GGATCGTGTC ATGACCCTAA

721 TTCGCCACTT TATTTCTTCT TTTACCAAGA TACTCCTCAT TGGACTTTAA TACACAGGAC
    AAGCGGTGAA ATAAAGAAGA AAATGGTTCT ATGAGGAGTA ACCTGAAATT ATGTGTCCTG

781 TAGTCTAAGG TATCACCAGG TAGTCCACTC CTGCTCGGAA TTCTTGACCC TCTTTCGGGA
    ATCAGATTCC ATAGTGGTCC ATCAGGTGAG GACGAGCCTT AAGAACTGGG AGAAAGCCCT

841 TTTAGAAGAA TAGGGCATGG ACCAGATGGG TTTAAACAAA TTCAATATCT TCCACTAGCT
    AAATCTTCTT ATCCCGTACC TGGTCTACCC AAATTTGTTT AAGTTATAGA AGGTGATCGA

901 TCACCTTGGG GTTGTTAAAA GATTTTTGAA CCACACACTG TGCTCATAAC AATCTTCATC
    AGTGGAACCC CAACAATTTT CTAAAAACTT GGTGTGTGAC ACGAGTATTG TTAGAAGTAG

961 TCTTAAAAGG ATTTTATTCT TCCTGGTATT GCCCTCACTC TCATCCCTGT ATTCCGTGCT
    AGAATTTTCC TAAAATAAGA AGGACCATAA CGGGAGTGAG AGTAGGGACA TAAGGCACGA
```

FIGURE 46C

```
1021 CAGTGGCTGA CACAGAAGAG TTCTTTATTG ATGTCCGCCC CCCACCCACT AGGATTCTCT
     GTCACCGACT GTGTCTTCTC AAGAAATAAC TACAGGCGGG GGGTGGGTGA TCCTAAGAGA

1081 GCTCTCCCCT CCCCCTACAG GCCTCCATCC TCTTCATCCT GTTCATTTTT CAGATCTCAG
     CGAGAGGGGA GGGGGATGTC CGGAGGTAGG AGAAGTAGGA CAAGTAAAAA GTCTAGAGTC

1141 TTCAAGCATC TCGTCCTCAG TGTGGTGTTT CCTGATCCCT CACTCTAATC CAAGTCTTTC
     AAGTTCGTAG AGCAGGAGTC ACACCACAAA GGACTAGGGA GTGAGATTAG GTTCAGAAAG

1201 TGTTTTATGC ACAGGTGGAA TCTTATTTCC GTTTGCGTCC AATCATGTAT TTTAATATGC
     ACAAAATACG TGTCCACCTT AGAATAAAGG CAAACGCAGG TTAGTACATA AAATTATACG

1261 ATGTATATAT GTATGTGCAT TTGTATGCAT GCGATTAAGA ACTAGAATAA TTAATAATTG
     TACATATATA CATACACGTA AACATACGTA CGCTAATTCT TGATCTTATT AATTATTAAC

1321 GAAAGCTCCA TGAAAGCTGG TTGGGGACTA ATTTTGTAAC TACTTTATTC CCAGATCCTG
     CTTTCGAGGT ACTTTCGACC AACCCCTGAT TAAAACATTG ATGAAATAAG GGTCTAGGAC

1381 TAATTTCTCT AAATAAACCC TGGAATCTTG CCTTATCTCC TTCAGGTTAA AAGCCAACTG
     ATTAAAGAGA TTTATTTGGG ACCTTAGAAC GGAATAGAGG AAGTCCAATT TTCGGTTGAC

1441 CAAGGTCTAA TGACTGCAGG ATCTAGCTAT CCATTGTTTC TGGCCGCCTA TGCGTGCACT
     GTTCCAGATT ACTGACGTCC TAGATCGATA GGTAACAAAG ACCGGCGGAT ACGCACGTGA

1501 GGGTGTCTGG CAGAGAGGCT GGGTAAATTG TAGTTTCATT GTAGCTGTCT GACTTGGATT
     CCCACAGACC GTCTCTCCGA CCCATTTAAC ATCAAAGTAA CATCGACAGA CTGAACCTAA

1561 TCTCACGCCT ACTTCACTGG AAACGCAAAC TCTCACAGCA TTTTGTTTTA GTTTCAGAAT
     AGAGTGCGGA TGAAGTGACC TTTGCGTTTG AGAGTGTCGT AAAACAAAAT CAAAGTCTTA

1621 CAGAGCAAAT TAGAAGTCTG AATTTCCTTC AACACTTGGA AATAATTTAT TTATTTGAAA
     GTCTCGTTTA ATCTTCAGAC TTAAAGGAAG TTGTGAACCT TTATTAAATA AATAAACTTT

1681 TATATTCATA ATTAATTCGT TATAAAAATG TATTAAATGC TTATTTGAGT CAGCAGAGGA
     ATATAAGTAT TAATTAAGCA ATATTTTTAC ATAATTTACG AATAAACTCA GTCGTCTCCT
```

FIGURE 46D

```
1741 AGATAGAAAC TTTATGAAAG TAGAAGGTGG ATCTCCTTTT TGCCTTCATT TTCAGAACAT
     TCTATCTTTG AAATACTTTC ATCTTCCACC TAGAGGAAAA ACGGAAGTAA AAGTCTTGTA

1801 CTCGTTTACA CCCATTAGTT GAAACATTAA TGTCATTTTA TTTTCGTCCT GATTATCTCA
     GAGCAAATGT GGGTAATCAA CTTTGTAATT ACAGTAAAAT AAAAGCAGGA CTAATAGAGT

1861 TAAAACATTT CTTAGAATAA CAGCAATACC TATCATTGAA GTTGGATAAG AAATATTTTG
     ATTTTGTAAA GAATCTTATT GTCGTTATGG ATAGTAACTT CAACCTATTC TTTATAAAAC

1921 CAATTGGTTT GCAACTTAAA AATCTGTTTG CATGACTCTT TTTCAGTGAA AGTAGGCAAG
     GTTAACCAAA CGTTGAATTT TTAGACAAAC GTACTGAGAA AAAGTCACTT TCATCCGTTC

1981 AGAAATTAAA ATTCAGAAAT ATCTCACCTA ATGTCAGAGG TAATATTGAT AATTTGTGTT
     TCTTTAATTT TAAGTCTTTA TAGAGTGGAT TACAGTCTCC ATTATAACTA TTAAACACAA

2041 TTACAAATAA TACATACAAC AATAATGAAA AATAAGTCCT ATCTATAGGC TCGTATCTCA
     AATGTTTATT ATGTATGTTG TTATTACTTT TTATTCAGGA TAGATATCCG AGCATAGAGT

2101 TGCCTATTTT TGGATGTATT TTTCA
     ACGGATAAAA ACCTACATAA AAAGT
```

FIGURE 47A

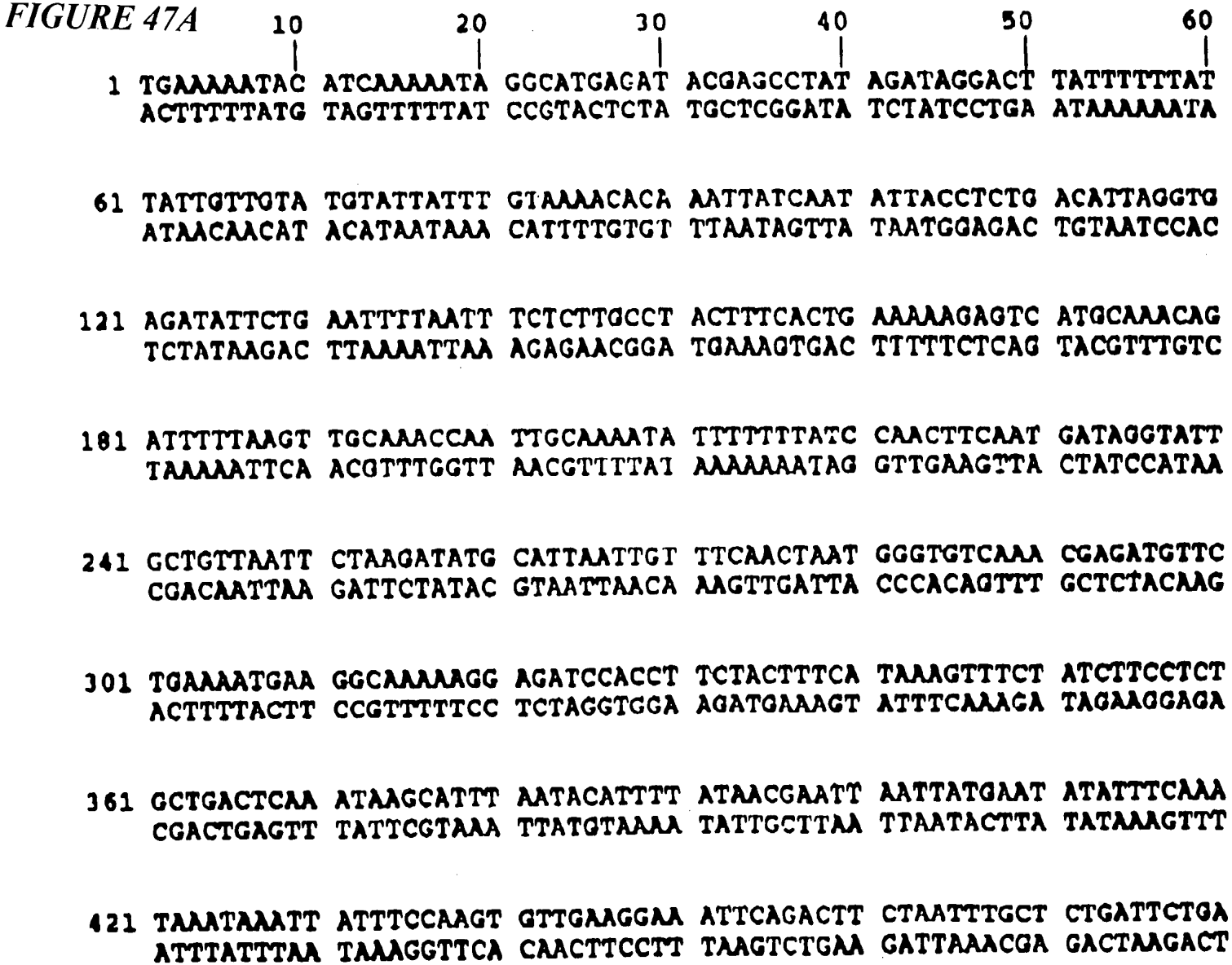

10 20 30 40 50 60
1 TGAAAAATAC ATCAAAAATA GGCATGAGAT ACGAGCCTAT AGATAGGACT TATTTTTTAT
ACTTTTTATG TAGTTTTTAT CCGTACTCTA TGCTCGGATA TCTATCCTGA ATAAAAAATA
61 TATTGTTGTA TGTATTATTT GTAAAACACA AATTATCAAT ATTACCTCTG ACATTAGGTG
ATAACAACAT ACATAATAAA CATTTTGTGT TTAATAGTTA TAATGGAGAC TGTAATCCAC
121 AGATATTCTG AATTTTAATT TCTCTTGCCT ACTTTCACTG AAAAAGAGTC ATGCAAACAG
TCTATAAGAC TTAAAATTAA AGAGAACGGA TGAAAGTGAC TTTTTCTCAG TACGTTTGTC
181 ATTTTTAAGT TGCAAACCAA TTGCAAAATA TTTTTTTATC CAACTTCAAT GATAGGTATT
TAAAAATTCA ACGTTTGGTT AACGTTTTAT AAAAAAATAG GTTGAAGTTA CTATCCATAA
241 GCTGTTAATT CTAAGATATG CATTAATTGT TTCAACTAAT GGGTGTCAAA CGAGATGTTC
CGACAATTAA GATTCTATAC GTAATTAACA AAGTTGATTA CCCACAGTTT GCTCTACAAG
301 TGAAAATGAA GGCAAAAAGG AGATCCACCT TCTACTTTCA TAAAGTTTCT ATCTTCCTCT
ACTTTTACTT CCGTTTTTCC TCTAGGTGGA AGATGAAAGT ATTTCAAAGA TAGAAGGAGA
361 GCTGACTCAA ATAAGCATTT AATACATTTT ATAACGAATT AATTATGAAT ATATTTCAAA
CGACTGAGTT TATTCGTAAA TTATGTAAAA TATTGCTTAA TTAATACTTA TATAAAGTTT
421 TAAATAAATT ATTTCCAAGT GTTGAAGGAA ATTCAGACTT CTAATTTGCT CTGATTCTGA
ATTTATTTAA TAAAGGTTCA CAACTTCCTT TAAGTCTGAA GATTAAACGA GACTAAGACT

FIGURE 47B

```
481 AACTAAAACA AATGCTCTGT GAGAGTTTGC GTTTCCAGTG AAGTAGCGTG AGAAATCCAA
    TTGATTTTGT TTACGAGACA CTCTCAAACG CAAAGGTCAC TTCATCGCAC TCTTTAGGTT

541 GTCAGACAGC TACATGAAAC TACATTTACC AGCTCTCTGC CAGACACCAG TGCACGATAG
    CAGTCTGTCG ATGTACTTTG ATGTAAATGG TCGAGAGACG GTCTGTGGTC ACGTGCTATC

601 CGCAGAACAT GTAGCTAGAT CTCAGTCATA GCTNNNNNNN NNNNNNNNNN AGACCTTGCA
    GCGTCTTGTA CATCGATCTA GAGTCAGTAT CGANNNNNNN NNNNNNNNNN TCTGGAACGT

661 GTTGGCTTTT AACCTGAAGG AGATAAGGCA AGATTCCAGG GTTTATTTAG AGAAATTACA
    CAACCGAAAA TTGGACTTCC TCTATTCCGT TCTAAGGTCC CAAATAAATC TCTTTAATGT

721 GGATCTGGGA ATAAAGTAGT TACAAAATTA GTCCCCAACC AGCTTTCATG GAGCTTTCAA
    CCTAGACCCT TATTTCATCA ATGTTTTAAT CAGGGGTTGG TCGAAAGTAC CTCGAAAGTT
```

FIGURE 47C

```
781 TTATTAATTA TTCTAGTTCT TAATCGCATG CATACAATGC ACATACATAT ATACATGCAT
    AATAATTAAT AAGATCAAGA ATTAGCGTAC GTATGTTACG TGTATGTATA TATGTACGTA

841 ATTAAAATAC ATGATTGGAC GCAAACGGAA ATAAGATTCC ACCTGTGCAT AAAACAGAAA
    TAATTTTATG TACTAACCTG CGTTTGCCTT TATTCTAAGG TGGACACGTA TTTTGTCTTT

901 GACTTGGTTA GAGTGAGGGA TCAGGAAACA CCACACTGAG GACGAGATGN NNNNNNNNNN
    CTGAACCAAT CTCACTCCCT AGTCCTTTGT GGTGTGACTC CTGCTCTACN NNNNNNNNNN

961 NTAGTGGGTG GGGGGCGGAC ATCAATAAAG AACTCTTCTG TGTCAGCCAC TGAGCACGGA
    NATCACCCAC CCCCCGCCTG TAGTTATTTC TTGAGAAGAC ACAGTCGGTG ACTCGTGCCT

1021 ATAAAGGGAT GAGAGTGAGG GCAANTACCA GAAGAATAAA ATCCTTTTAA GAGATGAAGA
     TATTTCCCTA CTCTCACTCC CGTTNATGGT CTTCTTATTT TAGGAAAATT CTCTACTTCT

1081 TTGTTATGAG CACAGTGTGT GGNTTCAAAA ATCTTTTAAC AACCCCAAGG TGAAGCTAGT
     AACAATACTC GTGTCACACA CCNAAGTTTT TAGAAAATTG TTGGGGTTCC ACTTCGATCA

1141 TGGAAGATAT TTGAATTTGT TTAAACCCAT CTGGTCCTAG CCCTATTCTT TGAATCCCGA
     ACCTTCTATA AACTTAAACA AATTTGGGTA GACCAGGATC GGGATAAGAA ACTTAGGGCT
```

FIGURE 47D

```
1201 AAGAGGGTCA AGAATTCCGA GCAGGAGTGG ACTACCTGGT GATACCTTAG ACTAGTCCTG
     TTCTCCCAGT TCTTAAGGCT CGTCCTCACC TGATGGACCA CTATGGAATC TGATCAGGAC

1261 TGTATTAAAG TCCAATGAGG AGTATCTTGG TAAAATAATA AATAAAGTCC CGAAAATCCC
     ACATAATTTC AGGTTACTCC TCATAGAACC ATTTTATTAT TTATTTCAGG GCTTTTAGGG

1321 AGTACTGTGC TAGGAGATTT ACATGCTATA TTATTTACTA TNNNNNNNNT AATTTGCAGA
     TCATGACACG ATCCTCTAAA TGTACGATAT AATAAATGAT ANNNNNNNNA TTAAACGTCT

1381 TAATATTATC CTCATCATAA AATAGGGTAA CTAACGCTGA GAGGGACTCG GTAACTTGTT
     ATTATAATAG GAGTAGTATT TTATCCCATT GATTGCGACT CTCCCTGAGC CATTGAACAA

1441 CAAGGCCACT AAGAAGTGGC AAAGTCAAAA CTGGAATTTT AATAAAAGAG TCTAGCTTGC
     GTTCCGGTGA TTCTTCACCG TTTCAGTTTT GACCTTAAAA TTATTTTCTC AGATCGAACG

1501 CTGTGTGGTT CTGCTTTTCT TAGAAAGTTG GANNAAGTCT CANATCAGTA CCCAGGAAAA
     GACACACCAA GACGAAAAGA ATCTTTCAAC CTNNTTCAGA GTNTAGTCAT GGGTCCTTTT

1561 ACAGCAAAAG ACCCGCTGGT AAAGACCTGT CCAGATTGCT GACCTGGTTC ACACANNTCC
```

FIGURE 47E

```
    TGTCGTTTTC TGGGCGACCA TTTCTGGACA GGTCTAACGA CTGGACCAAG TGTGTTNAGG

1621 AAGCTTGCCT CTGTTACTTC CAAGGAAGAA AGAATGCACA GAGAGGTAAA AAAACAAACA
     TTCGAACGGA GACAATGAAG GTTCCTTCTT TCTTACGTGT CTCTCCATTT TTTTGTTTGT

1681 AACCAAACAA AACAAAACAA AACAAAACAA AACAAAACAA AAGCAAAAAA AAACTTCCTC
     TTGGTTTGTT TTGTTTTGTT TTGTTTTGTT TTGTTTTGTT TTCGTTTTTT TTTGAAGGAG

1741 TGTCTTGCAG GGCTCCAGCA CTTGGAACCT TCCTACGTCC TANTTTCAGG TTCTCTCAGT
     ACAGAACGTC CCGAGGTCGT GAACCTTGGA AGGATGCAGG ATNAAAGTCC AAGAGAGTCA

1801 TCTACCCTCA ACCTGAGTGA CTGTCCTACC AGCAGCTTGT CGAGAACTCA GCCCTGCACC
     AGATGGGAGT TGGACTCACT GACAGGATGG TCGTCGAACA GCTCTTGAGT CGGGACGTGG

1861 GTTCCCAGCT ACCCTCCTCC TAACTCGAGG GGTGCT
     CAAGGGTCGA TGGGAGGAGG ATTGAGCTCC CCACGA
```

FIGURE 48A

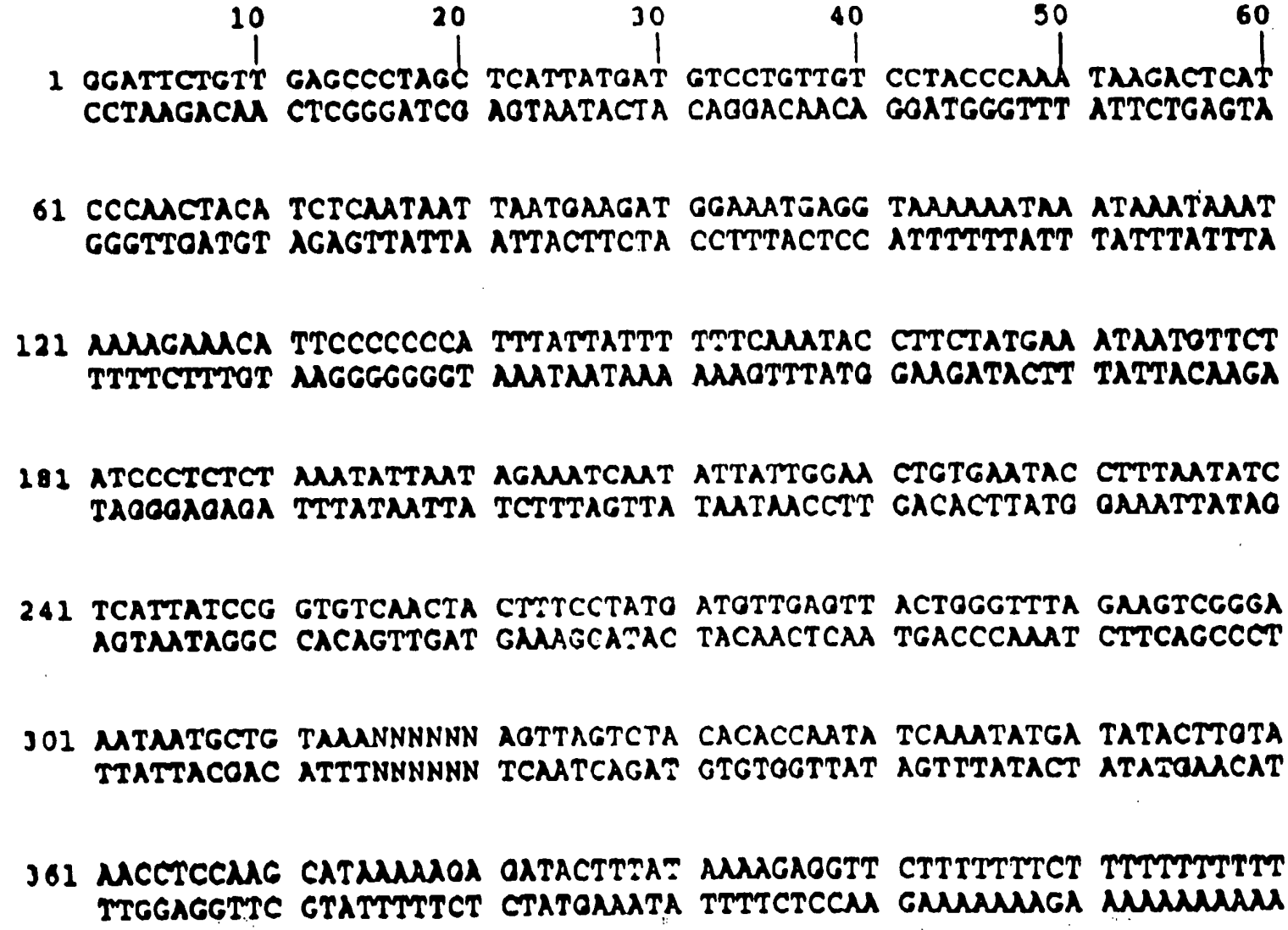

10 20 30 40 50 60
1 GGATTCTGTT GAGCCCTAGC TCATTATGAT GTCCTGTTGT CCTACCCAAA TAAGACTCAT
CCTAAGACAA CTCGGGATCG AGTAATACTA CAGGACAACA GGATGGGTTT ATTCTGAGTA
61 CCCAACTACA TCTCAATAAT TAATGAAGAT GGAAATGAGG TAAAAAATAA ATAAATAAAT
GGGTTGATGT AGAGTTATTA ATTACTTCTA CCTTTACTCC ATTTTTTATT TATTTATTTA
121 AAAAGAAACA TTCCCCCCCA TTTATTATTT TTTCAAATAC CTTCTATGAA ATAATGTTCT
TTTTCTTTGT AAGGGGGGGT AAATAATAAA AAAGTTTATG GAAGATACTT TATTACAAGA
181 ATCCCTCTCT AAATATTAAT AGAAATCAAT ATTATTGGAA CTGTGAATAC CTTTAATATC
TAGGGAGAGA TTTATAATTA TCTTTAGTTA TAATAACCTT GACACTTATG GAAATTATAG
241 TCATTATCCG GTGTCAACTA CTTTCCTATG ATGTTGAGTT ACTGGGTTTA GAAGTCGGGA
AGTAATAGGC CACAGTTGAT GAAAGCATAC TACAACTCAA TGACCCAAAT CTTCAGCCCT
301 AATAATGCTG TAAANNNNNN AGTTAGTCTA CACACCAATA TCAAATATGA TATACTTGTA
TTATTACGAC ATTTNNNNNN TCAATCAGAT GTGTGGTTAT AGTTTATACT ATATGAACAT
361 AACCTCCAAG CATAAAAAGA GATACTTTAT AAAAGAGGTT CTTTTTTTCT TTTTTTTTTT
TTGGAGGTTC GTATTTTTCT CTATGAAATA TTTTCTCCAA GAAAAAAAGA AAAAAAAAAA

FIGURE 48B

```
421 TCCAGATGGA GTTTCACTCC TGTCAGGCAG GCNGAGTGCA GTGGTGCCAT CTCGGCTCAC
    AGGTCTACCT CAAAGTGAGG ACAGTCCGTC CGNCTCACGT CACCACGGTA GAGCCGAGTG

481 TGCAACCTCC ACCTCCCATG TTCAAGGGAT TCTCCTTCCT CAGTCTCCTG AGTAGCTGGG
    ACGTTGGAGG TGGAGGGTAC AAGTTCCCTA AGAGGAAGGA GTCAGAGGAC TCATCGACCC

541 ATTACAGGTG TGCACCACCA CACCCAGCTA ATTTTTGTAT TTTTAATAGA GACAGGGTTT
    TAATGTCCAC ACGTGGTGGT GTGGGTCGAT TAAAAACATA AAAATTATCT CTGTCCCAAA

601 CATCGATGTT GGCCAGGCTA GTCTCGAACT CCTGACCTCT AGGTGATCCA CCCGCCTCAG
    GTAGCTACAA CCGGTCCGAT CAGAGCTTGA GGACTGGAGA TCCACTAGGT GGGCGGAGTC

661 CCTCCCAAAG TTGTAGAATT ACACGTGTGA GGCACTGCTC TGGCCAGGAG ATACATTTTT
    GGAGGGTTTC AACATCTTAA TGTGCACACT CCGTGACGAG ACCGGTCCTC TATGTAAAAA

721 GATAGGTTTA ATTTATAAAG ACACTGCACA GATTTGGACT TGCTGGGAAA TCACGATCCA
    CTATCCAAAT TAAATATTTC TGTGACGTGT CTAAACCTCA ACGACCCTTT AGTGCTAGGT
```

FIGURE 48C

```
781 GTATGCATTT GACCCAGCAA TTTTTATTGG TACTTAATGA TTATATCTCA ATTGATCAGG
    CATACGTAAA CTGGGTCGTT AAAAATAACC ATGAATTACT AATATAGAGT TAACTAGTCC

841 TTGAACTCTG TGCGAAGAAT TTGTGTGTGG ACATTTGAGA GGACAGTTTG GAGGCAAGGT
    AACTTGAGAC ACGCTTCTTA AACACACACC TGTAAACTCT CCTGTCAAAC CTCCGTTCCA

901 ATTTTAGTAG ATTTAAAGAA TTTGAATCTT GTTTGCAAGT TGGGGCATAT ACTGAGAAAG
    TAAAATCATC TAAATTTCTT AAACTTAGAA CAAACGTTCA ACCCCGTATA TGACTCTTTC

961 AGAAGACAAT GCAGATAAAT TGATATATTT ATTATGATGT ATGTTCAATA TGAAAGATCA
    TCTTCTGTTA CGTCTATTTA ACTATATAAA TAATACTACA TACAAGTTAT ACTTTCTAGT

1021 CAAAATATAA CATACATNNA TCTTACTTAA CATACCTCAG TTTTAGAGCT ACCGTATGTA
     GTTTTATATT GTATGTANNT AGAATGAATT GTATGGAGTC AAAATCTCGA TGGCATACAT

1081 GAAGAGTCCA TTTCTATTTA GGTAAGTTCC TTTAGTCCTT TTATTACTGG GCACTCTTAA
     CTTCTCAGGT AAAGATAAAT CCATTCAAGG AAATCAGGAA AATAATGACC CGTGAGAATT

1141 TTACATGTAG CTTGAAATAT GTCCAGTTTG AGCAGTGAAC TGAAAATGTC ATGTGATTAA
     AATGTACATC GAACTTTATA CAGGTCAAAC TCGTCACTTG ACTTTTACAG TACACTAATT

1201 GTACATATAT AATTTTTTTT CATAGTAGGT CAATAACCTC CTTTTATTGA CTAATGAATC
     CATGTATATA TTAAAAAAAA GTATCATCCA GTTATTGGAG GAAAATAACT GATTACTTAG

1261 AGTTCTCTAA TGATTATACG
     TCAAGAGATT ACTAATATGC
```

FIGURE 49A

```
                10          20          30          40          50          60
                |           |           |           |           |           |
  1 AATCAAAATA AAACAGTTAA AGTTTGATTA CTATAATCAA ACACAAAAAA AATGAATATT
    TTAGTTTTAT TTTGTCAATT TCAAACTAAT GATATTAGTT TGTGTTTTTT TTACTTATAA

 61 ATCTTTTATG TCAGTAGAGG GTGAATGAAT CCTTCAGGAT TTTGATGATA GTATCAGATA
    TAGAAAATAC AGTCATCTCC CACTTACTTA GGAAGTCCTA AAACTACTAT CATAGTCTAT

121 CCCAGCACTA TGCTAGAAGT TGTGAAGAAT TCACGAGATG AATAAATCAC AGATTCTGTC
    GGGTCGTGAT ACGATCTTCA ACACTTCTTA AGTGCTCTAC TTATTTAGTG TCTAAGACAG

181 CTCAAAATGG TTAGATCTAT TCAGGAAACA AAGCTAAAAA AACCCCACCA ATAACTAAAA
    GAGTTTTACC AATCTAGATA AGTCCTTTGT TTCGATTTTT TTGGGGTGGT TATTGATTTT

241 ATCAACCAAA TGAAAAACAA CAATCATAAA ATAAGTAAGT ACCTATAGAA AGAAAAGCTC
    TAGTTGGTTT ACTTTTTGTT GTTAGTATTT TATTCATTCA TGGATATCTT TCTTTTCGAG

301 AGAGGAGGTA AAAAGATAAC TCTTCCAAAA GGAATACTAT ATACTGTAAA CTGTGTACTG
    TCTCCTCCAT TTTTCTATTG AGAAGGTTTT CCTTATGATA TATGACATTT GACACATGAC

361 ATAGAAGGAA GAATTAGAAA NNNNNNNNTG TAAGTGGCAT ACATACTAAG CTAGTGTGAA
    TATCTTCCTT CTTAATCTTT NNNNNNNNAC ATTCACCGTA TGTATGATTC GATCACACTT
```

FIGURE 49B

```
421 CACAAGCCTA AATATGTAGT TGCTTCACAG AAGGTTAGAA GTAAATTAAC CTCATGAATT
    GTGTTCGGAT TTATACATCA ACGAAGTGTC TTCCAATCTT CATTTAATTG GAGTACTTAA

481 TCTTGAGAGA ACTTGTAAGG ACTAAGCTTT CGATTTTGGA GAAAGATTTT AATACCAAAT
    AGAACTCTCT TGAACATTCC TGATTCGAAA GCTAAAACCT CTTTCTAAAA TTATGGTTTA

541 AAAAAGTACC TTTGTTTGGT AATCTCAATC ATTATAATAG TGCTTAGATA ATACCTAGGA
    TTTTTCATGG AAACAAACCA TTAGAGTTAG TAATATTATC ACGAATCTAT TATGGATCCT

601 ACAAATTAAA TATTAAATTT ACTTTAAAAA AAAGTACATG ATTGGGGAAT CACAACTGGC
    TGTTTAATTT ATAATTTAAA TGAAATTTTT TTTCATGTAC TAACCCCTTA GTGTTGACCG

661 CTTACTAGAT TCTCTNNNNN NATATGCACT GAAAAGAATG AAAAACACTG AACCAAATAT
    GAATGATCTA AGAGANNNNN NTATACGTGA CTTTTCTTAC TTTTTGTGAC TTGGTTTATA

721 NTGTTTTTTT AAGTTTAAAA TTAAATTGGA AAAAAATAGT AAGGAATATC AGAAGCAAAA
    NACAAAAAAA TTCAAATTTT AATTTAACCT TTTTTTATCA TTCCTTATAG TCTTCGTTTT
```

FIGURE 49C

```
781 AAATAAAATG AAAGCAAGAA TCCTCAGAGG TAGCACGAAA TTTGGCTTTG CTTAGATGGA
    TTTATTTTAC TTTCGTTCTT AGGAGTCTCC ATCGTGCTTT AAACCGAAAC GAATCTACCT

841 TCTATCAAAG CTATGGCCCA TGAAAAGGAT TCAGGAGTTA GTTTAAAGCT GGTTCACATA
    AGATAGTTTC GATACCGGGT ACTTTTCCTA AGTCCTCAAT CAAATTTCGA CCAAGTGTAT

901 ATGGAATCTA GCAGAAGACT GTGCATAAAG GTGGTCTAAG AACAACAATA TCCTGACCAG
    TACCTTAGAT CGTCTTCTGA CACGTATTTC CACCAGATTC TTGTTGTTAT AGGACTGGTC

961 GTGAGGGGGC TCACNCTNAA TNCCAGCACT TTGGGAGCCC AAGGTGGGTG GATCACGAGG
    CACTCCCCCG AGTGNGANTT ANGGTCGTGA AACCCTCGGG TTCCACCCAC CTAGTGCTCC

1021 TCAGGAGTTT GAGACCAGCC TGACCAACAT GGTGAAACCG CGTCTCTACT AAAAATAGAA
     AGTCCTCAAA CTCTGGTCGG ACTGGTTGTA CCACTTTGGC GCAGAGATGA TTTTTATCTT

1081 AAATTAGCCG NGCCTACGTG CTTCTAATCC CAGCTGAACT CAGGAGACTG AGACAGGAGA
     TTTAATCGGC NCGGATGCAC GAAGATTAGG GTCGACTTGA GTCCTCTGAC TCTGTCCTCT

1141 ATCACTTGAA CCCAGCATGC AAGCTTNNNN NNGCCACTGC ACTCCAGCCT AGGGTGCAAA
     TAGTGAACTT GGGTCGTACG TTCGAANNNN NNCGGTGACG TGAGGTCGGA TCCCACGTTT

1201 AAAAAAAAAA ANGACACATT ACTCAGGTAA GGTAATCAAT AA
     TTTTTTTTTT TNCTGTGTAA TGAGTCCATT CCATTAGTTA TT
```

FIGURE 50A

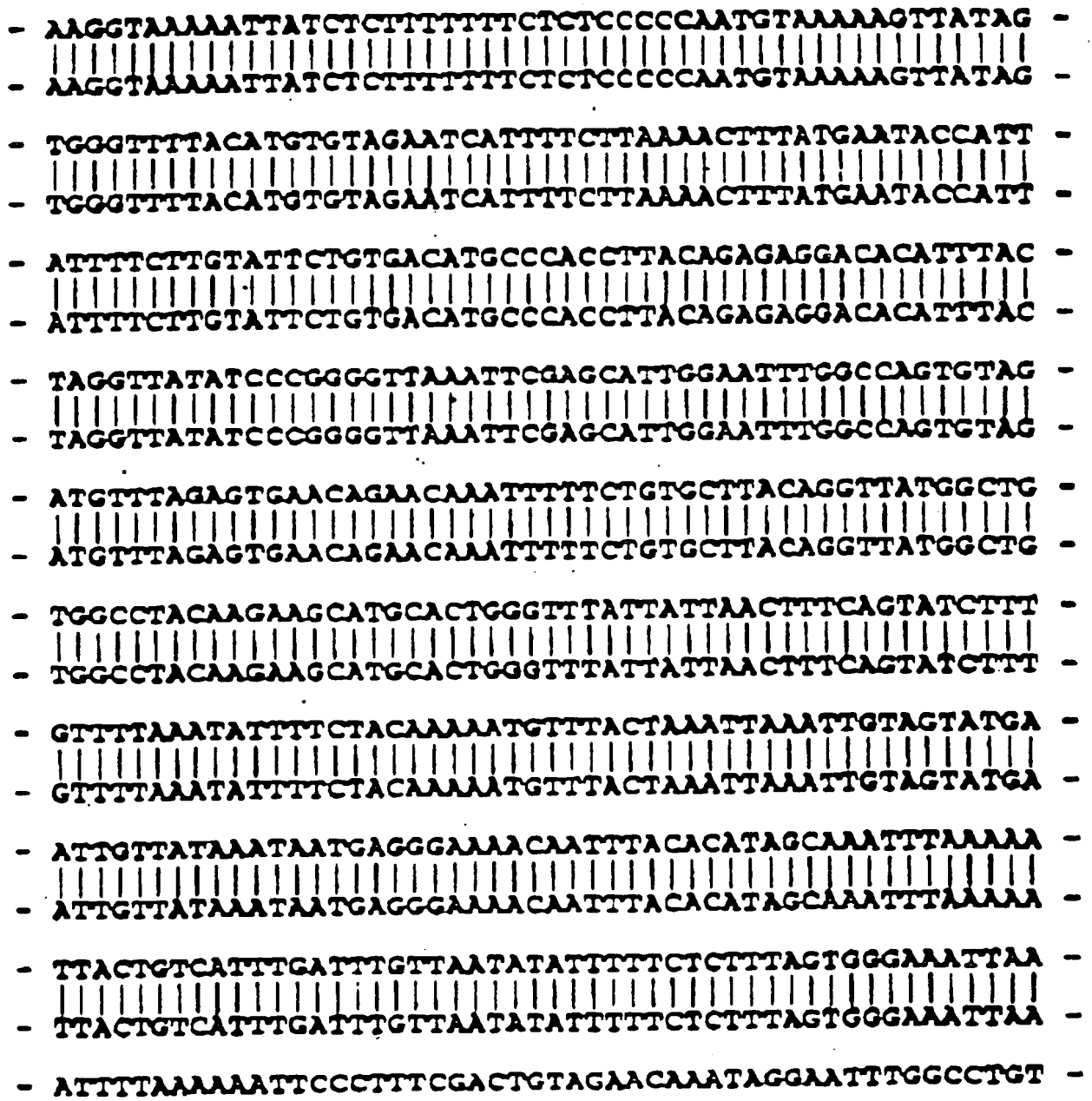

- AAGGTAAAAATTATCTCTTTTTTTCTCTCCCCCAATGTAAAAAGTTATAG -
- AAGGTAAAAATTATCTCTTTTTTTCTCTCCCCCAATGTAAAAAGTTATAG -
- TGGGTTTTACATGTGTAGAATCATTTTCTTAAAACTTTATGAATACCATT -
- TGGGTTTTACATGTGTAGAATCATTTTCTTAAAACTTTATGAATACCATT -
- ATTTTCTTGTATTCTGTGACATGCCCACCTTACAGAGAGGACACATTTAC -
- ATTTTCTTGTATTCTGTGACATGCCCACCTTACAGAGAGGACACATTTAC -
- TAGGTTATATCCCGGGGTTAAATTCGAGCATTGGAATTTGGCCAGTGTAG -
- TAGGTTATATCCCGGGGTTAAATTCGAGCATTGGAATTTGGCCAGTGTAG -
- ATGTTTAGAGTGAACAGAACAAATTTTTCTGTGCTTACAGGTTATGGCTG -
- ATGTTTAGAGTGAACAGAACAAATTTTTCTGTGCTTACAGGTTATGGCTG -
- TGGCCTACAAGAAGCATGCACTGGGTTTATTATTAACTTTCAGTATCTTT -
- TGGCCTACAAGAAGCATGCACTGGGTTTATTATTAACTTTCAGTATCTTT -
- GTTTTAAATATTTTCTACAAAAATGTTTACTAAATTAAATTGTAGTATGA -
- GTTTTAAATATTTTCTACAAAAATGTTTACTAAATTAAATTGTAGTATGA -
- ATTGTTATAAATAATGAGGGAAAACAATTTACACATAGCAAATTTAAAAA -
- ATTGTTATAAATAATGAGGGAAAACAATTTACACATAGCAAATTTAAAAA -
- TTACTGTCATTTGATTTGTTAATATATTTTTCTCTTTAGTGGGAAATTAA -
- TTACTGTCATTTGATTTGTTAATATATTTTTCTCTTTAGTGGGAAATTAA -
- ATTTTAAAAAATTCCCTTTCGACTGTAGAACAAATAGGAATTTGGCCTGT -

FIGURE 50B

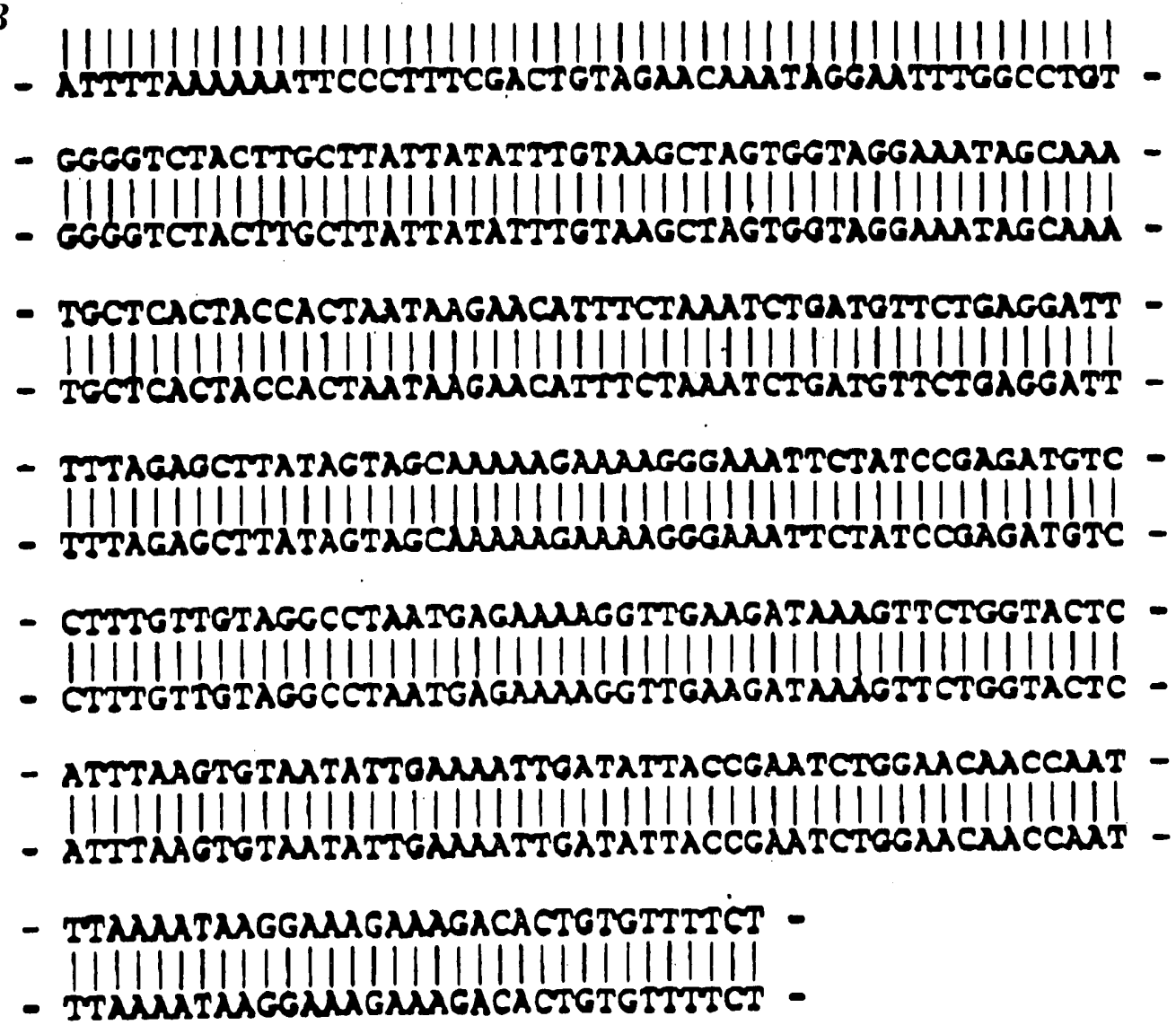

- ATTTTAAAAAATTCCCTTTCGACTGTAGAACAAATAGGAATTTGGCCTGT -
- GGGGTCTACTTGCTTATTATATTTGTAAGCTAGTGGTAGGAAATAGCAAA -
- GGGGTCTACTTGCTTATTATATTTGTAAGCTAGTGGTAGGAAATAGCAAA -
- TGCTCACTACCACTAATAAGAACATTTCTAAATCTGATGTTCTGAGGATT -
- TGCTCACTACCACTAATAAGAACATTTCTAAATCTGATGTTCTGAGGATT -
- TTTAGAGCTTATAGTAGCAAAAAGAAAAGGGAAATTCTATCCGAGATGTC -
- TTTAGAGCTTATAGTAGCAAAAAGAAAAGGGAAATTCTATCCGAGATGTC -
- CTTTGTTGTAGGCCTAATGAGAAAAGGTTGAAGATAAAGTTCTGGTACTC -
- CTTTGTTGTAGGCCTAATGAGAAAAGGTTGAAGATAAAGTTCTGGTACTC -
- ATTTAAGTGTAATATTGAAAATTGATATTACCGAATCTGGAACAACCAAT -
- ATTTAAGTGTAATATTGAAAATTGATATTACCGAATCTGGAACAACCAAT -
- TTAAAATAAGGAAAGAAAGACACTGTGTTTTCT -
- TTAAAATAAGGAAAGAAAGACACTGTGTTTTCT -

FIGURE 51A

```
                10         20         30         40         50         60
      1 AGAAAACACA GTGTCTTTCT TTCCTTATTT TAAATTGGTT GTTCCAGATT CGGTAATATC
        TCTTTTGTGT CACAGAAAGA AAGGAATAAA ATTTAACCAA CAAGGTCTAA GCCATTATAG

     61 AATTTTCAAT ATTACACTTA AATGAGTACC AGAACTTTAT CTTCAACCTT TTCTCATTAG
        TTAAAAGTTA TAATGTGAAT TTACTCATGG TCTTGAAATA GAAGTTGGAA AAGAGTAATC

    121 GCCTACAACA AAGGACATCT CGGATAGAAT TTCCCTTTTC TTTTTGCTAC TATAAGCTCT
        CGGATGTTGT TTCCTGTAGA GCCTATCTTA AAGGGAAAAG AAAAACGATG ATATTCGAGA

    181 AAAAATCCTC AGAACATCAG ATTTAGAAAT GTTCTTATTA GTGGTAGTGA GCATTTGCTA
        TTTTTAGGAG TCTTGTAGTC TAAATCTTTA CAAGAATAAT CACCATCACT CGTAAACGAT

    241 TTTCCTACCA CTAGCTTACA AATATAATAA GCAAGTAGAC CCCACAGGCC AAATTCCTAT
        AAAGGATGGT GATCGAATGT TTATATTATT CGTTCATCTG GGGTGTCCGG TTTAAGGATA

    301 TTGTTCTACA GTCGAAAGGG AATTTTTTAA AATTTAATTT CCCACTAAAG AGAAAAATAT
        AACAAGATGT CAGCTTTCCC TTAAAAAATT TTAAATTAAA GGGTGATTTC TCTTTTTATA

    361 ATTAACAAAT CAAATGACAG TAATTTTTAA ATTTGCTATG TGTAAATTGT TTTCCCTCAT
        TAATTGTTTA GTTTACTGTC ATTAAAAATT TAAACGATAC ACATTTAACA AAAGGGAGTA

    421 TATTTATAAC AATTCATACT ACAATTTAAT TTAGTAAACA TTTTTGTAGA AAATATTTAA
        ATAAATATTG TTAAGTATGA TGTTAAATTA AATCATTTGT AAAAACATCT TTTATAAATT
```

FIGURE 51B

```
481 AACAAAGATA CTGAAAGTTA ATATNAAACC CAGTGCATGC TTCTTGTAGG CCACAGCCAT
    TTGTTTCTAT GACTTTCAAT TATANTTTGG GTCACGTACG AAGAACATCC GGTGTCGGTA

541 AACCTGTAAG CACAGAAAAA TTTGTTCTGT TACTCTAAAC ATCTACACTG GCCAAATTCC
    TTGGACATTC GTGTCTTTTT AAACAAGACA ATGAGATTTG TAGATGTGAC CGGTTTAAGG

601 AATGCTCGAA TTTAACCCCG GGATATAACC TAGTAAATGT GTCCTCTCTG TAAGGTGGGC
    TTACGAGCTT AAATTGGGGC CCTATATTGG ATCATTTACA CAGGAGAGAC ATTCCACCCG

661 ATGTCACAGA ATACAAGAAA ATAATGGTAT TCATAAAGTT TTAAGAAAAT GATTCTACAC
    TACAGTGTCT TATGTTCTTT TATTACCATA AGTATTTCAA AATTCTTTTA CTAAGATGTG

721 ATGTAAAACC CACTATAACT TTTTACATTG GGGGAGAGAA AAAAAGAGAT AATTTTTACC
    TACATTTTGG GTGATATTGA AAAATGTAAC CCCCTCTCTT TTTTTCTCTA TTAAAAATGG

781 TT
    AA
```

FIGURE 52A

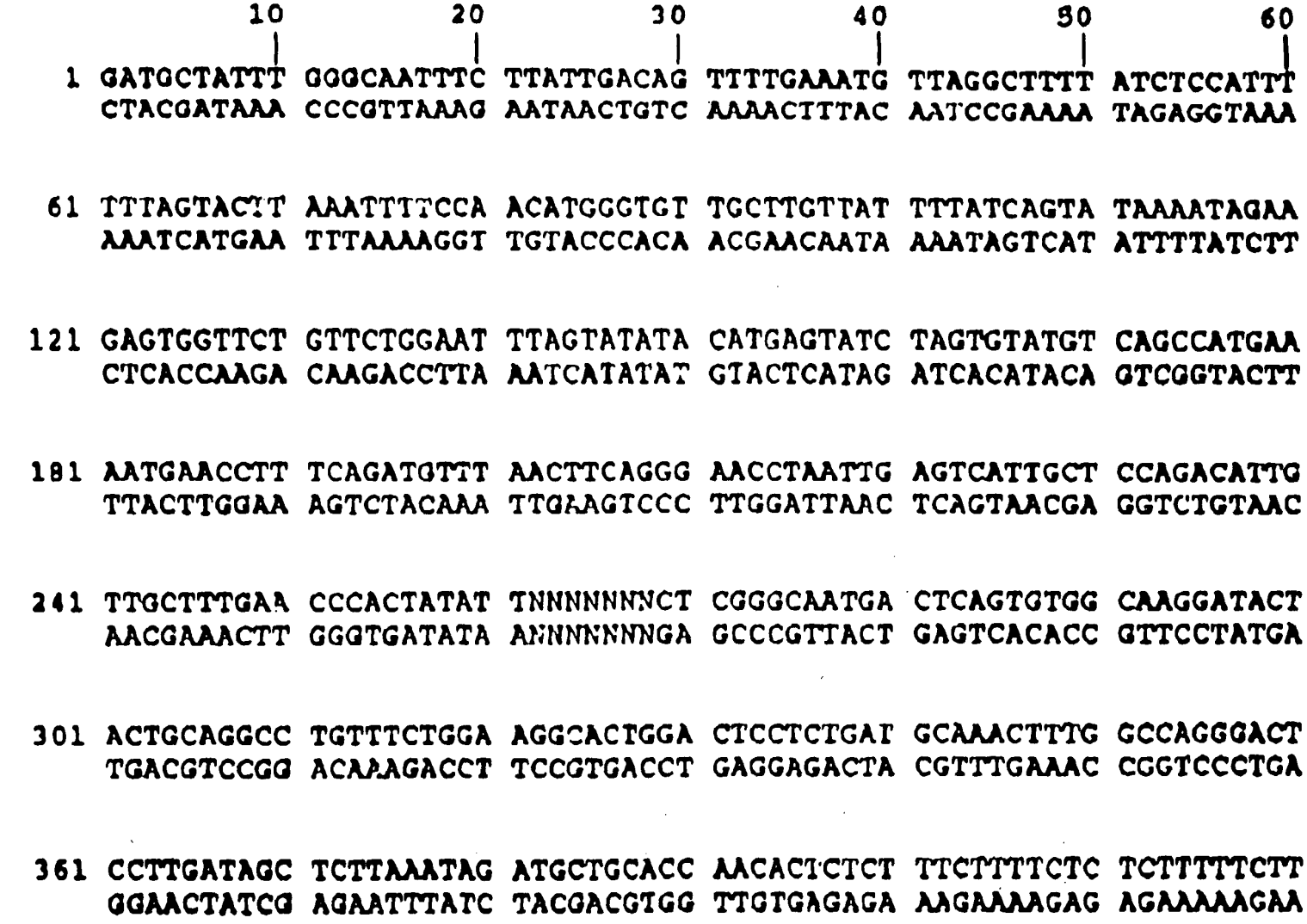

10 20 30 40 50 60
1 GATGCTATTT GGGCAATTTC TTATTGACAG TTTTGAAATG TTAGGCTTTT ATCTCCATTT
CTACGATAAA CCCGTTAAAG AATAACTGTC AAAACTTTAC AATCCGAAAA TAGAGGTAAA
61 TTTAGTACTT AAATTTTCCA ACATGGGTGT TGCTTGTTAT TTTATCAGTA TAAAATAGAA
AAATCATGAA TTTAAAAGGT TGTACCCACA ACGAACAATA AAATAGTCAT ATTTTATCTT
121 GAGTGGTTCT GTTCTGGAAT TTAGTATATA CATGAGTATC TAGTGTATGT CAGCCATGAA
CTCACCAAGA CAAGACCTTA AATCATATAT GTACTCATAG ATCACATACA GTCGGTACTT
181 AATGAACCTT TCAGATGTTT AACTTCAGGG AACCTAATTG AGTCATTGCT CCAGACATTG
TTACTTGGAA AGTCTACAAA TTGAAGTCCC TTGGATTAAC TCAGTAACGA GGTCTGTAAC
241 TTGCTTTGAA CCCACTATAT TNNNNNNNCT CGGGCAATGA CTCAGTGTGG CAAGGATACT
AACGAAACTT GGGTGATATA ANNNNNNNGA GCCCGTTACT GAGTCACACC GTTCCTATGA
301 ACTGCAGGCC TGTTTCTGGA AGGCACTGGA CTCCTCTGAT GCAAACTTTG GCCAGGGACT
TGACGTCCGG ACAAAGACCT TCCGTGACCT GAGGAGACTA CGTTTGAAAC CGGTCCCTGA
361 CCTTGATAGC TCTTAAATAG ATGCTGCACC AACACTCTCT TTCTTTTCTC TCTTTTTCTT
GGAACTATCG AGAATTTATC TACGACGTGG TTGTGAGAGA AAGAAAAGAG AGAAAAAGAA

FIGURE 52B

```
421 TATTCAATAT TAGACTACAA GCAGTCTAAG GACTTCTCAG GGTTTCTAGC TCTCTCTCAT
    ATAAGTTATA ATCTGATGTT CGTCAGATTC CTGAAGAGTC CCAAAGATCG AGAGAGAGTA

481 TTCACACATG CTTTCCTAGT AATCTCTACT CATATATCTT ACTGCTACGC TGGGGCCAGA
    AAGTGTGTAC GAAAGGATCA TTAGAGATGA GTATATAGAA TGACGATGCG ACCCCGGTCT

541 TAACNNNNNN CTTCCATTTT GTTTTTATCT CTATTCTTCT TCCCCTTCTG CTTTCATTAT
    ATTGNNNNNN GAAGGTAAAA CAAAAATAGA GATAAGAAGA AGGGGAAGAC GAAAGTAATA

601 TGAAACTTTC TGCTTTCATT ATTGAAACTT TCCCAGATTT GTTCTGCTTA ACCTGGCATT
    ACTTTGAAAG ACGAAAGTAA TAACTTTGAA AGGGTCTAAA CAAGACGAAT TGGACCGTAA

661 GGAACTGTTT CCTCTTCCCT GTGCTGCTTT CTCCCATTGC CATGTCCTTT TTTTTTTTTT
    CCTTGACAAA GGAGAAGGGA CACGACGAAA GAGGGTAACG GTACAGGAAA AAAAAAAAAA

721 TTTTTTTTTT TGAGACAGTG TCACTCTGTT GCCCAGGCTG GAGTGCAATG GTGCAATCTT
    AAAAAAAAAA ACTCTGTCAC AGTGAGACAA CGGGTCCGAC CTCACGTTAC CACGTTAGAA
```

FIGURE 52C

```
781 GGCCACTGCA ACCCCCGCCT CCCGGGTTCA AGTGATTCTC CTGCCTCAGC CTCCTGAGTA
    CCGGTGACGT TGGGGGCGGA GGGCCCAAGT TCACTAAGAG GACGGAGTCG GAGGACTCAT

841 GCTGGGATTA CAGGTGCCCA CCACTATGCC CGGCTGATTT TTGTATTTTT AGTAGAGATN
    CGACCCTAAT GTCCACGGGT GGTGATACGG GCCGACTAAA AACATAAAAA TCATCTCTAN

901 NNNNNNNTTT CACCATNGCT GATCAGGCTG GTCTCGAACT CCTGACCGCA GTGANTCCGC
    NNNNNNNAAA GTGGTANCGA CTAGTCCGAC CAGAGCTTGA GGACTGGCGT CACTNAGGCG

961 CCTCCTTGGC CTCCCAAAGT GCTGAGATTA CAGGCATGAG TCACTGCGNC CAGCCACCAT
    GGAGGAACCG GAGGGTTTCA CGACTCTAAT GTCCGTACTC AGTGACGCNG GTCGGTGGTA

1021 TATTCTCTAG AGGTGAGAGA ACACTGGCTC TTCTAACAAG TTGAAATTTG ATAGAGACC
     ATAAGAGATC TCCACTCTCT TGTGACCGAG AAGATTGTTC AACTTTAAAC TATCTCTGG
```

FIGURE 53A

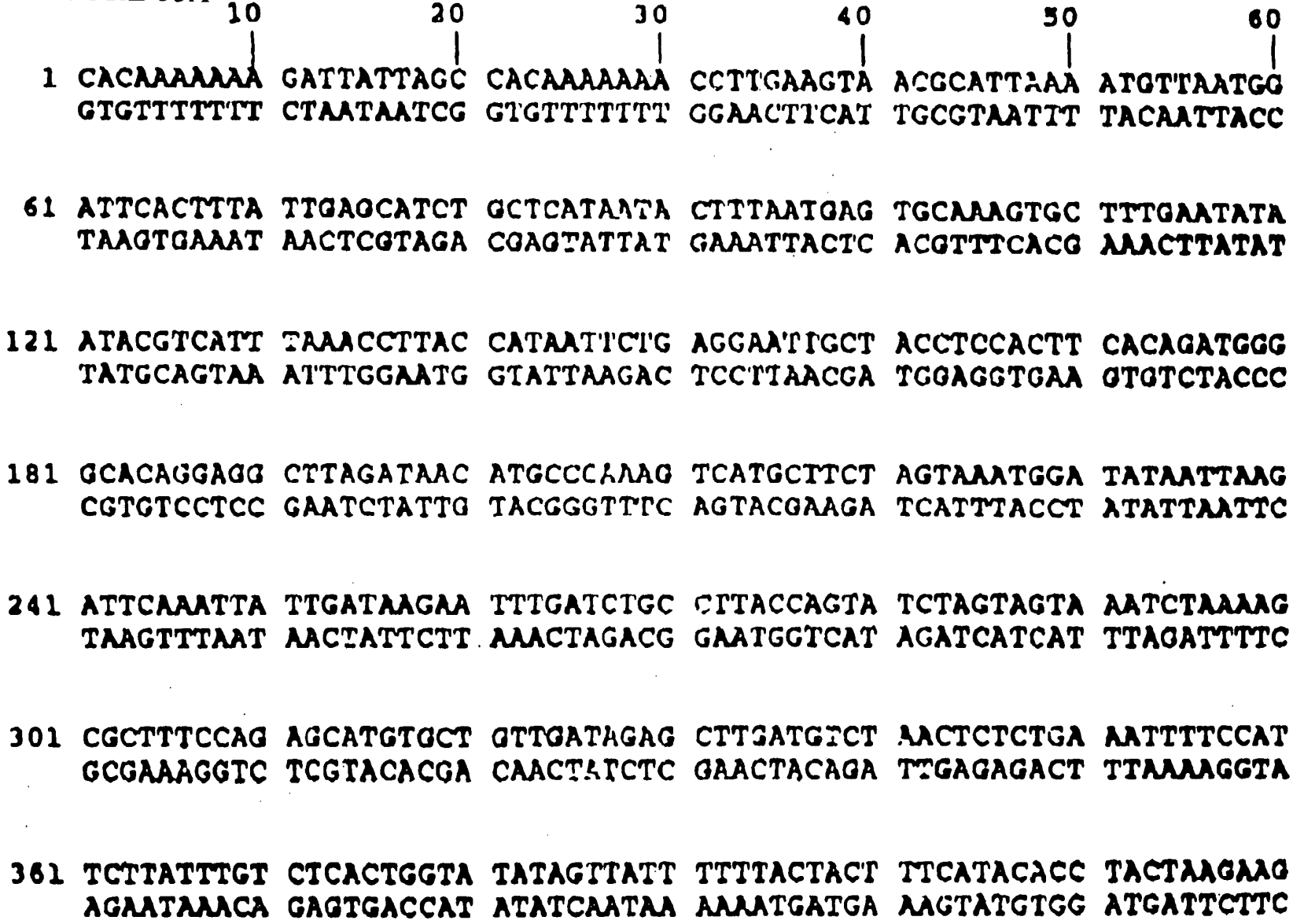

10 20 30 40 50 60
1 CACAAAAAAA GATTATTAGC CACAAAAAAA CCTTGAAGTA ACGCATTAAA ATGTTAATGG
GTGTTTTTTT CTAATAATCG GTGTTTTTTT GGAACTTCAT TGCGTAATTT TACAATTACC
61 ATTCACTTTA TTGAGCATCT GCTCATAATA CTTTAATGAG TGCAAAGTGC TTTGAATATA
TAAGTGAAAT AACTCGTAGA CGAGTATTAT GAAATTACTC ACGTTTCACG AAACTTATAT
121 ATACGTCATT TAAACCTTAC CATAATTCTG AGGAATTGCT ACCTCCACTT CACAGATGGG
TATGCAGTAA ATTTGGAATG GTATTAAGAC TCCTTAACGA TGGAGGTGAA GTGTCTACCC
181 GCACAGGAGG CTTAGATAAC ATGCCCAAAG TCATGCTTCT AGTAAATGGA TATAATTAAG
CGTGTCCTCC GAATCTATTG TACGGGTTTC AGTACGAAGA TCATTTACCT ATATTAATTC
241 ATTCAAATTA TTGATAAGAA TTTGATCTGC CTTACCAGTA TCTAGTAGTA AATCTAAAAG
TAAGTTTAAT AACTATTCTT AAACTAGACG GAATGGTCAT AGATCATCAT TTAGATTTTC
301 CGCTTTCCAG AGCATGTGCT GTTGATAGAG CTTGATGTCT AACTCTCTGA AATTTTCCAT
GCGAAAGGTC TCGTACACGA CAACTATCTC GAACTACAGA TTGAGAGACT TTAAAAGGTA
361 TCTTATTTGT CTCACTGGTA TATAGTTATT TTTTACTACT TTCATACACC TACTAAGAAG
AGAATAAACA GAGTGACCAT ATATCAATAA AAAATGATGA AAGTATGTGG ATGATTCTTC

FIGURE 53B

```
421 ACAGGAGGAT CAAAGATAGG ATTTCATTTA GAATGCCTAA AGCTTCACGT ATTTTAATTC
    TGTCCTCCTA GTTTCTATCC TAAAGTAAAT CTTACGGATT TCGAAGTGCA TAAAATTAAG

481 AGAATAAGAT TCAGGCAGAC CACCAGTATA TGCCATGGTC CCTGGTTATC TTTCAGCAGG
    TCTTATTCTA AGTCCGTCTG GTGGTCATAT ACGGTACCAG GGACCAATAG AAAGTCGTCC

541 TGACCGAGAA AGAAAACATG GTAATGTTTA TGAAATGGTG GGTTCTTGTA GTTTCACTTC
    ACTGGCTCTT TCTTTTGTAC CATTACAAAT ACTTTACCAC CCAAGAACAT CAAAGTGAAG

601 AACATATCTG CCTTTACTGT ATTAAGATGA TGGATTAACT TATTCTTGAT ATGGGCATGT
    TTGTATAGAC GGAAATGACA TAATTCTACT ACCTAATTGA ATAAGAACTA TACCCGTACA

661 AAAACAATAT ACTTTTACTA AACAGCTACA GAGAGACAAA TGTGTTTCCA GACAAACTTA
    TTTTGTTATA TGAAAATGAT TTGTCGATGT CTCTCTGTTT ACACAAAGGT CTGTTTGAAT

721 AGAGACTGAG TGTTCAAACT GAATAATCTC GACCTTAATT GTAACTATAT TTTATGAAAT
    TCTCTGACTC ACAAGTTTGA CTTATTAGAG CTGGAATTAA CATTGATATA AAATACTTTA
```

FIGURE 53C

```
781  CCAGCTGTAA GGCAAAAACA GACTTCTTTG GGCCTACCAC GGGCATTTTG TTCCTGTTAN
     GGTCGACATT CCGTTTTTGT CTGAAGAAAC CCGGATGGTG CCCGTAAAAC AAGGACAATN

841  NNNTACTCCA AACCTTAAAC CCACGTCCAC TTAAATAATG GCCTGGAAAT AAATGTCATT
     NNNATGAGGT TTGGAATTTG GGTGCAGGTG AATTTATTAC CGGACCTTTA TTTACAGTAA

901  ATCTGATATT ATACTGAGAT GTTTAGTTAT GAAATCAAAA GTGGAGAATT TCAATCTGTC
     TAGACTATAA TATGACTCTA CAAATCAATA CTTTAGTTTT CACCTCTTAA AGTTAGACAG

961  CTGTAAGCTT TCTCTGCGGT CACGACCCTC ATGCACTCAG GCTGTGCGGT GCAGCATGCT
     GACATTCGAA AGAGACGCCA GTGCTGGGAG TACGTGAGTC CGACACGCCA CGTCGTACGA

1021 CTGTCATGTC TGTTTTCTTC TGCCTGTACA CGGGTGGTTG TTCCTGTCTA CCTGTTTGAG
     GACAGTACAG ACAAAAGAAG ACGGACATGT GCCCACCAAC AAGGACAGAT GGACAAACTC

1081 GAAATATGAA TACGTNNNNN NCTAGAATCT ACTGCACATG CAATAAGGAA ACAATCAGTA
     CTTTATACTT ATGCANNNNN NGATCTTAGA TGACGTGTAC GTTATTCCTT TGTTAGTCAT

1141 AGAATCACTT TCTCGTGGAA AATTCATTAG AATTAACATC TCGTTTTAAA ATGCTCTATC
     TCTTAGTGAA AGAGCACCTT TTAAGTAATC TTAATTGTAG AGCAAAATTT TACGAGATAG
```

FIGURE 53D

```
1201 AAAGTGTAAA TAATTCCTCT CTCTTTTCCC TTTTTCACTA AGGAGTTTGT ATATTAAACA
     TTTCACATTT ATTAAGGAGA GAGAAAAGGG AAAAAGTGAT TCCTCAAACA TATAATTTGT

1261 GAATTTCAAG TAATGTATTA TAAATTTATT TAANNTATTT ACAATAAAAT GCCACGTATA
     CTTAAAGTTC ATTACATAAT ATTTAAATAA ATTNNATAAA TGTTATTTTA CGGTGCATAT

1321 AGCATCAAGC AACATGANNN NNNCATTGGT AGAAAGCACA ATACATAGTC AAAACAGCAG
     TCGTAGTTCG TTGTACTNNN NNNGTAACCA TCTTTCGTGT TATGTATCAG TTTTGTCGTC

1381 AGTATTAAAT AAACAGAAAA TTTGCAAAAG GCAAGTAAAG AATATACATA TACTTAATTA
     TCATAATTTA TTTGTCTTTT AAACGTTTTC CGTTCATTTC TTATATGTAT ATGAATTAAT

1441 TACATAAAAT ATTGATACAG GAGGTAGAAA GAAATTTAGT AAGCAGATAA TGGGGGCAAC
     ATGTATTTTA TAACTATGTC CTCCATCTTT CTTTAAATCA TTCGTCTATT ACCCCCGTTG

1501 AGAGTCCTCA GCAGAGCTTC CCTTCTAACA AAAAGCAGCC CAATAAATTA TTTTTTTTTT
     TCTCAGGAGT CGTCTCGAAG GGAAGATTGT TTTTCGTCGG GTTATTTAAT AAAAAAAAAA

1561 CTAACAAAAA GCAGCCTGAA AAATCGAGCT GCAAACATAG ATTAGCAATC GGCTGAAAGT
```

FIGURE 53E

```
     GATTGTTTTT CGTCGGACTT TTTAGCTCGA CGTTTGTATC TAATCGTTAG CCGACTTTCA

1621 GCGGGAGAAT GCTGGCAGCT GTGCCAATAG TAAAGGGCTA CCTGGAGCCG GGCGCGTGGC
     CGCCCTCTTA CGACCGTCGA CACGGTTATC ATTTCCCGAT GGACCTCGGC CCGCGCACCG

1681 TCACGCTGTA ATCCCAGCAC TTTGGGAGGG CGAGGCAACG CGGATCACCT GAGGTCGGGA
     AGTGCGACAT TAGGGTCGTG AAACCCTCCC GCTCCGTTGC GCCTAGTGGA CTCCAGCCCT

1741 GTTTGAGATC AGCCCGACCA ACATGGAGAA ACCCCGTCTC TACTAAAAAA AAAAAAAAAA
     CAAACTCTAG TCGGGCTGGT TGTACCTCTT TGGGGCAGAG ATGATTTTTT TTTTTTTTTT

1801 AAAGGCAAAA AATGAGCCGG GCATGGTGGC ACATGCCTTG CACATCCCAG CTGAGGCAGG
     TTTCCGTTTT TTACTCGGCC CGTACCACCG TGTACGGAAC GTGTAGGGTC GACTCCGTCC

1861 AGAATTCACT TGAACCTGGG AGGTAGAGAT TGCGGTGAAG CGAGATCACG TCATTGCACT
     TCTTAAGTGA ACTTGGACCC TCCATCTCTA ACGCCACTTC GCTCTAGTGC AGTAACGTGA

1921 CCAGCCTGGG CAAAAAGAGC AAAACTTAGT CTCAAAAAAA AAAANNCAAA GAAAAAA
     GGTCGGACCC GTTTTTCTCG TTTTGAATCA GAGTTTTTTT TTTTNNGTTT CTTTTTT
```

FIGURE 54

Genomic Organization of PSM Gene

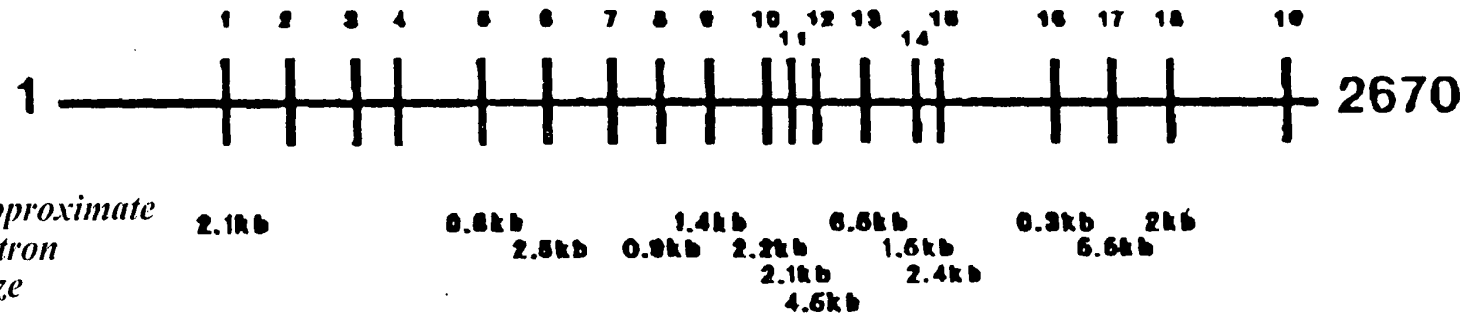


FIGURE 55A

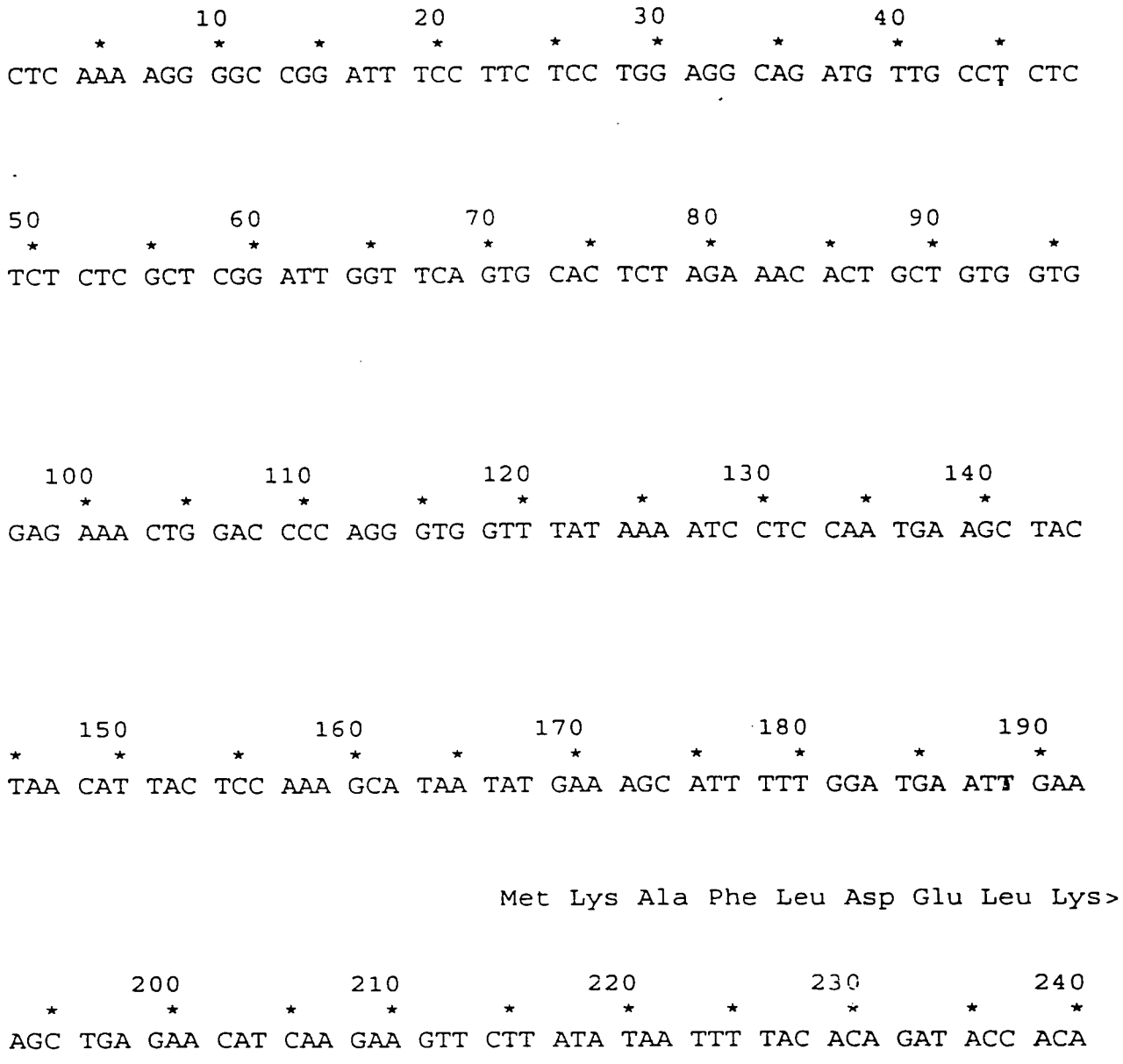

10 20 30 40
CTC AAA AGG GGC CGG ATT TCC TTC TCC TGG AGG CAG ATG TTG CCT CTC
50 60 70 80 90
TCT CTC GCT CGG ATT GGT TCA GTG CAC TCT AGA AAC ACT GCT GTG GTG
100 110 120 130 140
GAG AAA CTG GAC CCC AGG GTG GTT TAT AAA ATC CTC CAA TGA AGC TAC
150 160 170 180 190
TAA CAT TAC TCC AAA GCA TAA TAT GAA AGC ATT TTT GGA TGA ATT GAA
Met Lys Ala Phe Leu Asp Glu Leu Lys>
200 210 220 230 240
AGC TGA GAA CAT CAA GAA GTT CTT ATA TAA TTT TAC ACA GAT ACC ACA
Ala Glu Asn Ile Lys Lys Phe Leu Tyr Asn Phe Thr Gln Ile Pro His>

FIGURE 55B

```
          250          260          270          280
     *    *    *    *    *    *    *    *    *
TTT AGC AGG AAC AGA ACA AAA CTT TCA GCT TGC AAA GCA AAT TCA ATC

 Leu Ala Gly Thr Glu Gln Asn Phe Gln Leu Ala Lys Gln Ile Gln Ser>

290          300          310          320          330
  *    *    *    *    *    *    *    *    *    *
CCA GTG GAA AGA ATT TGG CCT GGA TTC TGT TGA GCT AGC ACA TTA TGA

 Gln Trp Lys Glu Phe Gly Leu Asp Ser Val Glu Leu Ala His Tyr Asp>

  340          350          360          370          380
    *    *    *    *    *    *    *    *    *
TGT CCT GTT GTC CTA CCC AAA TAA GAC TCA TCC CAA CTA CAT CTC AAT

 Val Leu Leu Ser Tyr Pro Asn Lys Thr His Pro Asn Tyr Ile Ser Ile>

    390          400          410          420          430
*    *    *    *    *    *    *    *    *    *
AAT TAA TGA AGA TGG AAA TGA GAT TTT CAA CAC ATC ATT ATT TGA ACC

 Ile Asn Glu Asp Gly Asn Glu Ile Phe Asn Thr Ser Leu Phe Glu Pro>

       440          450          460          470          480
  *    *    *    *    *    *    *    *    *    *
ACC TCC TCC AGG ATA TGA AAA TGT TTC GGA TAT TGT ACC ACC TTT CAG

 Pro Pro Pro Gly Tyr Glu Asn Val Ser Asp Ile Val Pro Pro Phe Ser>
```

FIGURE 55C

```
                490         500         510         520
            *   *   *   *   *   *   *   *   *   *
TGC TTT CTC TCC TCA AGG AAT GCC AGA GGG CGA TCT AGT GTA TGT TAA

Ala Phe Ser Pro Gln Gly Met Pro Glu Gly Asp Leu Val Tyr Val Asn>

530         540         550         560         570
  *   *   *   *   *   *   *   *   *   *
CTA TGC ACG AAC TGA AGA CTT CTT TAA ATT GGA ACG GGA CAT GAA AAT

Tyr Ala Arg Thr Glu Asp Phe Phe Lys Leu Glu Arg Asp Met Lys Ile>

  580         590         600         610         620
    *   *   *   *   *   *   *   *   *
CAA TTG CTC TGG GAA AAT TGT AAT TGC CAG ATA TGG GAA AGT TTT CAG

Asn Cys Ser Gly Lys Ile Val Ile Ala Arg Tyr Gly Lys Val Phe Arg>

    630         640         650         660         670
*   *   *   *   *   *   *   *   *   *   *
AGG AAA TAA GGT TAA AAA TGC CCA GCT GGC AGG GGC CAA AGG AGT CAT

Gly Asn Lys Val Lys Asn Ala Gln Leu Ala Gly Ala Lys Gly Val Ile>

      680         690         700         710         720
  *   *   *   *   *   *   *   *   *   *   *
TCT CTA CTC CGA CCC TGC TGA CTA CTT TGC TCC TGG GGT GAA GTC CTA

Leu Tyr Ser Asp Pro Ala Asp Tyr Phe Ala Pro Gly Val Lys Ser Tyr>
```

FIGURE 55D

```
          730         740         750         760
      *   *   *   *   *   *   *   *   *   *   *
TCC AGA TGG TTG GAA TCT TCC TGG AGG TGG TGT CCA GCG TGG AAA TAT

 Pro Asp Gly Trp Asn Leu Pro Gly Gly Gly Val Gln Arg Gly Asn Ile>

770         780         790         800         810
  *   *   *   *   *   *   *   *   *   *   *   *   *   *   *
CCT AAA TCT GAA TGG TGC AGG AGA CCC TCT CAC ACC AGG TTA CCC AGC

 Leu Asn Leu Asn Gly Ala Gly Asp Pro Leu Thr Pro Gly Tyr Pro Ala>

  820         830         840         850         860
    *   *   *   *   *   *   *   *   *   *   *   *   *   *   *
AAA TGA ATA TGC TTA TAG GCG TGG AAT TGC AGA GGC TGT TGG TCT TCC

 Asn Glu Tyr Ala Tyr Arg Arg Gly Ile Ala Glu Ala Val Gly Leu Pro>

    870         880         890         900         910
*   *   *   *   *   *   *   *   *   *   *   *   *   *   *
AAG TAT TCC TGT TCA TCC AAT TGG ATA CTA TGA TGC ACA GAA GCT CCT

 Ser Ile Pro Val His Pro Ile Gly Tyr Tyr Asp Ala Gln Lys Leu Leu>

       920         930         940         950         960
  *   *   *   *   *   *   *   *   *   *   *   *   *   *   *
AGA AAA AAT GGG TGG CTC AGC ACC ACC AGA TAG CAG CTG GAG AGG AAG

 Glu Lys Met Gly Gly Ser Ala Pro Pro Asp Ser Ser Trp Arg Gly Ser>
```

FIGURE 55E

```
          970         980         990        1000
     *    *    *    *    *    *    *    *    *
TCT CAA AGT GCC CTA CAA TGT TGG ACC TGG CTT TAC TGG AAA CTT TTC

 Leu Lys Val Pro Tyr Asn Val Gly Pro Gly Phe Thr Gly Asn Phe Ser>

1010        1020        1030        1040        1050
  *    *    *    *    *    *    *    *    *    *
TAC ACA AAA AGT CAA GAT GCA CAT CCA CTC TAC CAA TGA AGT GAC AAG

 Thr Gln Lys Val Lys Met His Ile His Ser Thr Asn Glu Val Thr Arg>

 1060        1070        1080        1090        1100
   *    *    *    *    *    *    *    *    *    *
AAT TTA CAA TGT GAT AGG TAC TCT CAG AGG AGC AGT GGA ACC AGA CAG

 Ile Tyr Asn Val Ile Gly Thr Leu Arg Gly Ala Val Glu Pro Asp Arg>

  1110        1120        1130        1140        1150
*    *    *    *    *    *    *    *    *    *    *
ATA TGT CAT TCT GGG AGG TCA CCG GGA CTC ATG GGT GTT TGG TGG TAT

 Tyr Val Ile Leu Gly Gly His Arg Asp Ser Trp Val Phe Gly Gly Ile>

     1160        1170        1180        1190        1200
  *    *    *    *    *    *    *    *    *    *    *
TGA CCC TCA GAG TGG AGC AGC TGT TGT TCA TGA AAT TGT GAG GAG CTT

 Asp Pro Gln Ser Gly Ala Ala Val Val His Glu Ile Val Arg Ser Phe>

         1210        1220        1230        1240
     *    *    *    *    *    *    *    *    *
TGG AAC ACT GAA AAA GGA AGG GTG GAG ACC TAG AAG AAC AAT TTT GTT

 Gly Thr Leu Lys Lys Glu Gly Trp Arg Pro Arg Arg Thr Ile Leu Phe>
```

FIGURE 55F

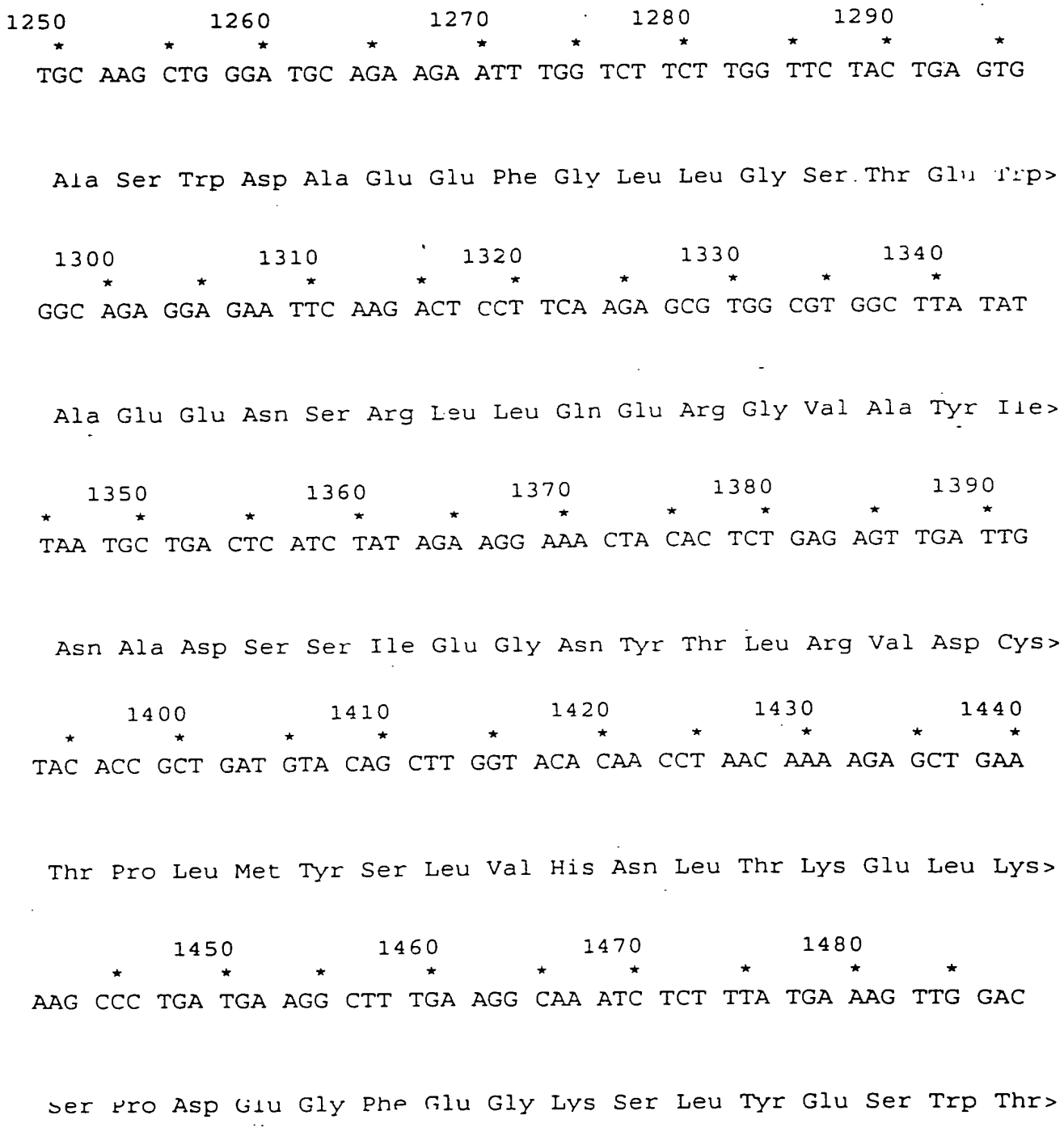

1250 1260 1270 1280 1290
TGC AAG CTG GGA TGC AGA AGA ATT TGG TCT TCT TGG TTC TAC TGA GTG
Ala Ser Trp Asp Ala Glu Glu Phe Gly Leu Leu Gly Ser Thr Glu Trp>
1300 1310 1320 1330 1340
GGC AGA GGA GAA TTC AAG ACT CCT TCA AGA GCG TGG CGT GGC TTA TAT
Ala Glu Glu Asn Ser Arg Leu Leu Gln Glu Arg Gly Val Ala Tyr Ile>
1350 1360 1370 1380 1390
TAA TGC TGA CTC ATC TAT AGA AGG AAA CTA CAC TCT GAG AGT TGA TTG
Asn Ala Asp Ser Ser Ile Glu Gly Asn Tyr Thr Leu Arg Val Asp Cys>
1400 1410 1420 1430 1440
TAC ACC GCT GAT GTA CAG CTT GGT ACA CAA CCT AAC AAA AGA GCT GAA
Thr Pro Leu Met Tyr Ser Leu Val His Asn Leu Thr Lys Glu Leu Lys>
1450 1460 1470 1480
AAG CCC TGA TGA AGG CTT TGA AGG CAA ATC TCT TTA TGA AAG TTG GAC
Ser Pro Asp Glu Gly Phe Glu Gly Lys Ser Leu Tyr Glu Ser Trp Thr>

FIGURE 55G

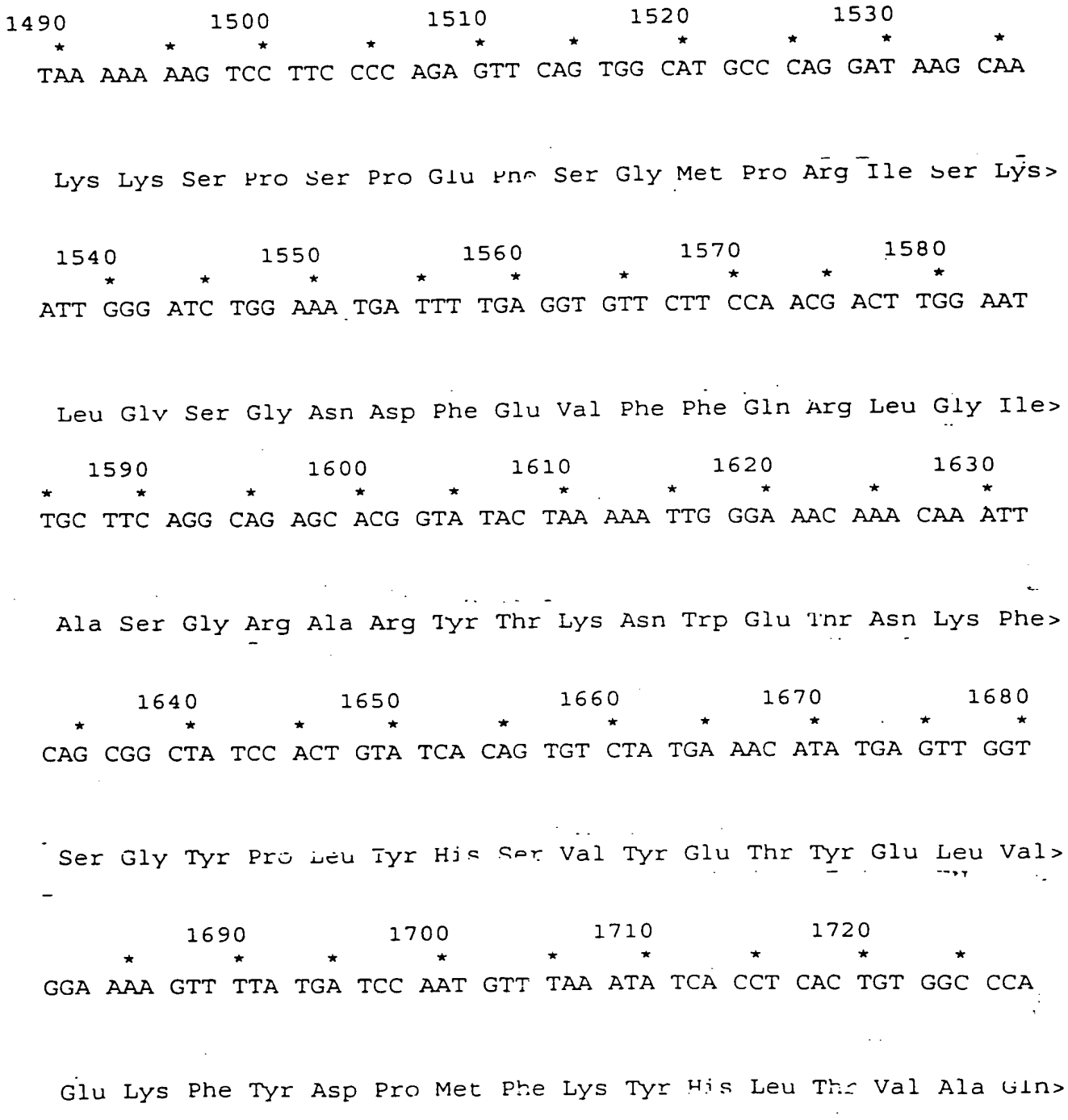

1490 1500 1510 1520 1530
TAA AAA AAG TCC TTC CCC AGA GTT CAG TGG CAT GCC CAG GAT AAG CAA
Lys Lys Ser Pro Ser Pro Glu Phe Ser Gly Met Pro Arg Ile Ser Lys>
1540 1550 1560 1570 1580
ATT GGG ATC TGG AAA TGA TTT TGA GGT GTT CTT CCA ACG ACT TGG AAT
Leu Glv Ser Gly Asn Asp Phe Glu Val Phe Phe Gln Arg Leu Gly Ile>
1590 1600 1610 1620 1630
TGC TTC AGG CAG AGC ACG GTA TAC TAA AAA TTG GGA AAC AAA CAA ATT
Ala Ser Gly Arg Ala Arg Tyr Thr Lys Asn Trp Glu Thr Asn Lys Phe>
1640 1650 1660 1670 1680
CAG CGG CTA TCC ACT GTA TCA CAG TGT CTA TGA AAC ATA TGA GTT GGT
Ser Gly Tyr Pro Leu Tyr His Ser Val Tyr Glu Thr Tyr Glu Leu Val>
1690 1700 1710 1720
GGA AAA GTT TTA TGA TCC AAT GTT TAA ATA TCA CCT CAC TGT GGC CCA
Glu Lys Phe Tyr Asp Pro Met Phe Lys Tyr His Leu Thr Val Ala Gln>

FIGURE 55H

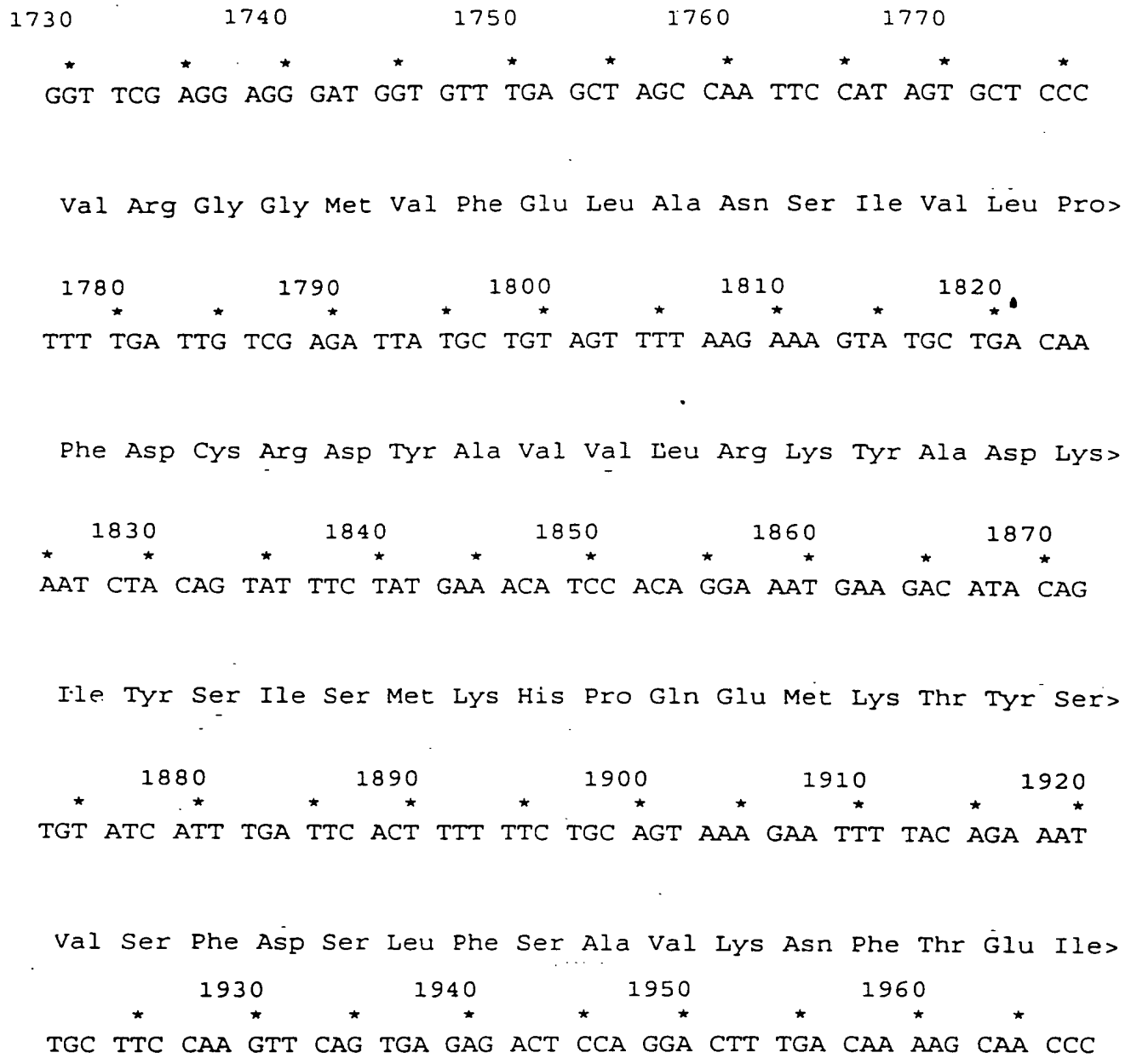
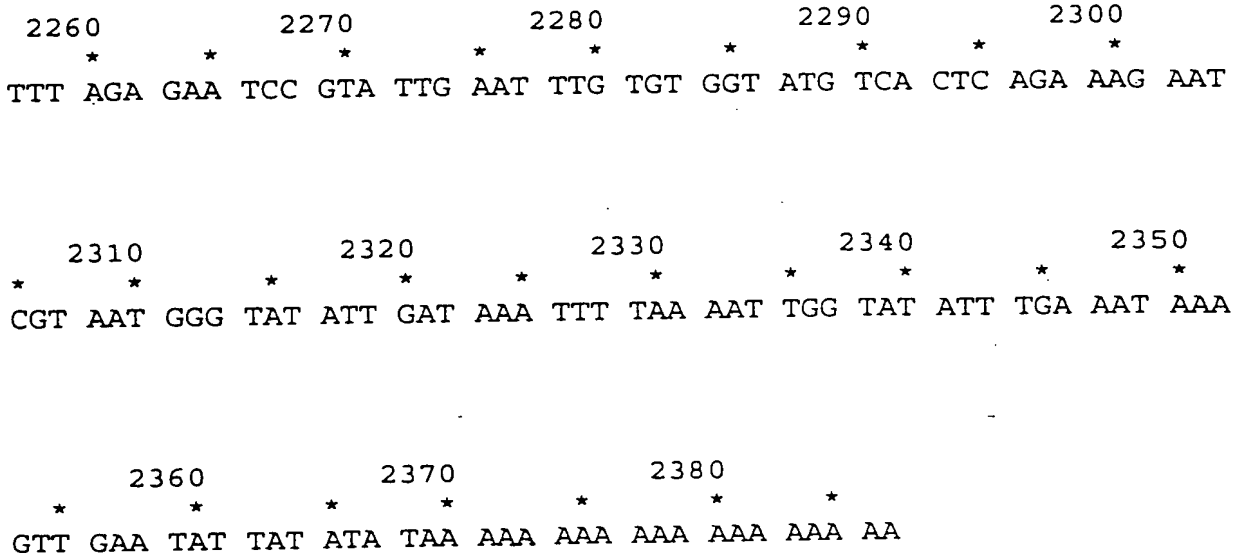

1730 1740 1750 1760 1770
GGT TCG AGG AGG GAT GGT GTT TGA GCT AGC CAA TTC CAT AGT GCT CCC
Val Arg Gly Gly Met Val Phe Glu Leu Ala Asn Ser Ile Val Leu Pro>
1780 1790 1800 1810 1820
TTT TGA TTG TCG AGA TTA TGC TGT AGT TTT AAG AAA GTA TGC TGA CAA
Phe Asp Cys Arg Asp Tyr Ala Val Val Leu Arg Lys Tyr Ala Asp Lys>
1830 1840 1850 1860 1870
AAT CTA CAG TAT TTC TAT GAA ACA TCC ACA GGA AAT GAA GAC ATA CAG
Ile Tyr Ser Ile Ser Met Lys His Pro Gln Glu Met Lys Thr Tyr Ser>
1880 1890 1900 1910 1920
TGT ATC ATT TGA TTC ACT TTT TTC TGC AGT AAA GAA TTT TAC AGA AAT
Val Ser Phe Asp Ser Leu Phe Ser Ala Val Lys Asn Phe Thr Glu Ile>
1930 1940 1950 1960
TGC TTC CAA GTT CAG TGA GAG ACT CCA GGA CTT TGA CAA AAG CAA CCC
Ala Ser Lys Phe Ser Glu Arg Leu Gln Asp Phe Asp Lys Ser Asn Pro>

FIGURE 55I

```
1970          1980          1990          2000          2010
   *     *     *     *     *     *     *     *     *     *
AAT AGT ATT AAG AAT GAT GAA TGA TCA ACT CAT GTT TCT GGA AAG AGC

 Ile Val Leu Arg Met Met Asn Asp Gln Leu Met Phe Leu Glu Arg Ala>

 2020          2030          2040          2050          2060
    *     *     *     *     *     *     *     *     *
ATT TAT TGA TCC ATT AGG GTT ACC AGA CAG GCC TTT TTA TAG GCA TGT

 Phe Ile Asp Pro Leu Gly Leu Pro Asp Arg Pro Phe Tyr Arg His Val>

   2070          2080          2090          2100          2110
*     *     *     *     *     *     *     *     *     *
CAT CTA TGC TCC AAG CAG CCA CAA CAA GTA TGC AGG GGA GTC ATT CCC

 Ile Tyr Ala Pro Ser Ser His Asn Lys Tyr Ala Gly Glu Ser Phe Pro>

     2120          2130          2140          2150          2160
  *     *     *     *     *     *     *     *     *     *
AGG AAT TTA TGA TGC TCT GTT TGA TAT TGA AAG CAA AGT GGA CCC TTC

 Gly Ile Tyr Asp Ala Leu Phe Asp Ile Glu Ser Lys Val Asp Pro Ser>

         2170          2180          2190          2200
     *     *     *     *     *     *     *     *     *
CAA GGC CTG GGG AGA AGT GAA GAG ACA GAT TTA TGT TGC AGC CTT CAC

 Lys Ala Trp Gly Glu Val Lys Arg Gln Ile Tyr Val Ala Ala Phe Thr>

2210          2220          2230          2240          2250
   *     *     *     *     *     *     *     *     *     *
AGT GCA GGC AGC TGC AGA GAC TTT GAG TGA AGT AGC CTA AGA GGA TTC

 Val Gln Ala Ala Ala Glu Thr Leu Ser Glu Val Ala
```

FIGURE 58J


2260
2270
2280
2290
2300
TTT AGA GAA TCC GTA TTG AAT TTG TGT GGT ATG TCA CTC AGA AAG AAT
2310
2320
2330
2340
2350
CGT AAT GGG TAT ATT GAT AAA TTT TAA AAT TGG TAT ATT TGA AAT AAA
2360
2370
2380
GTT GAA TAT TAT ATA TAA AAA AAA AAA AAA AAA AA

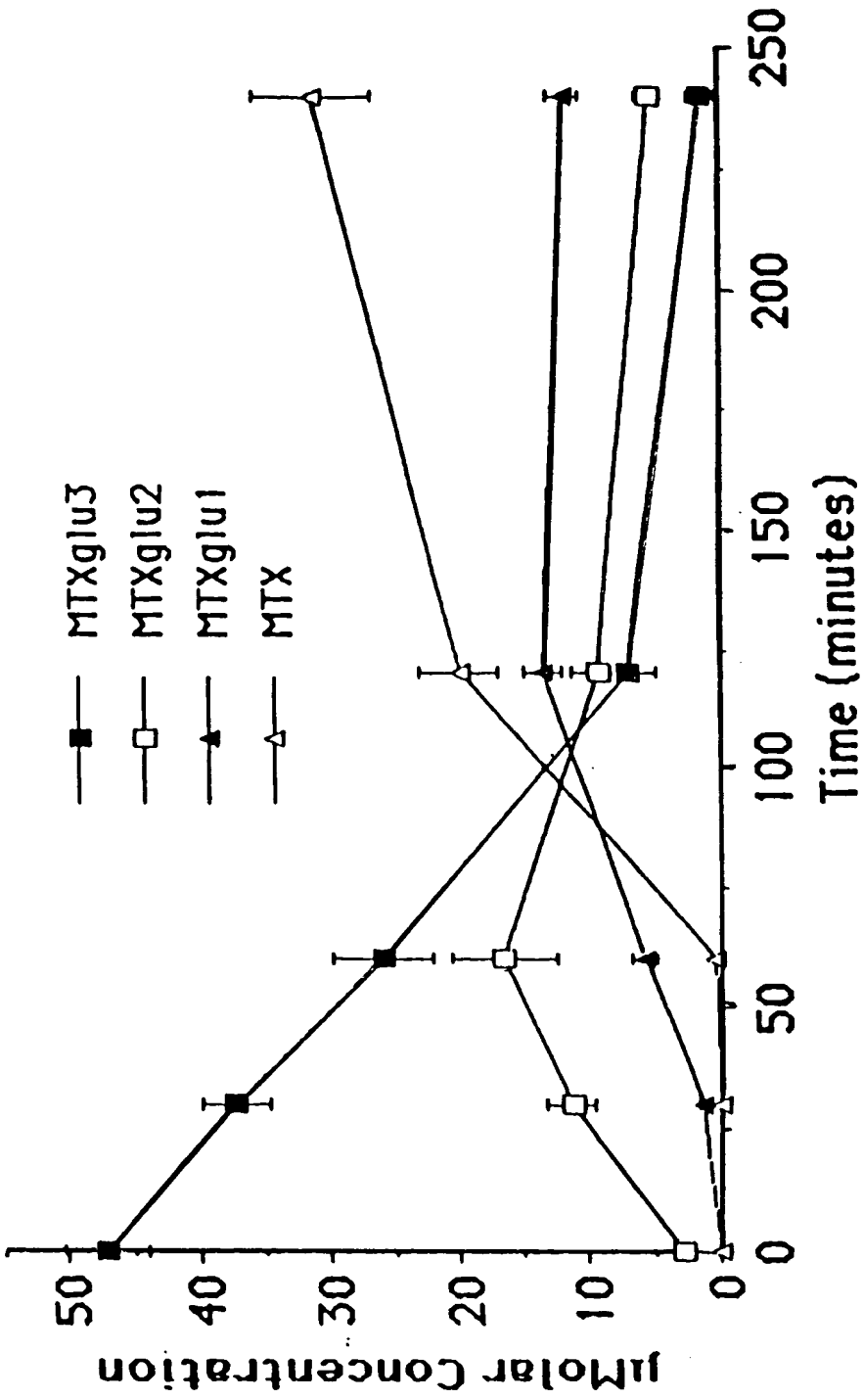

MTXglu3
MTXglu2
MTXglu1
MTX
μMolar Concentration
Time (minutes)
0
10
20
30
40
50
0
50
100
150
200
250

FIGURE 56

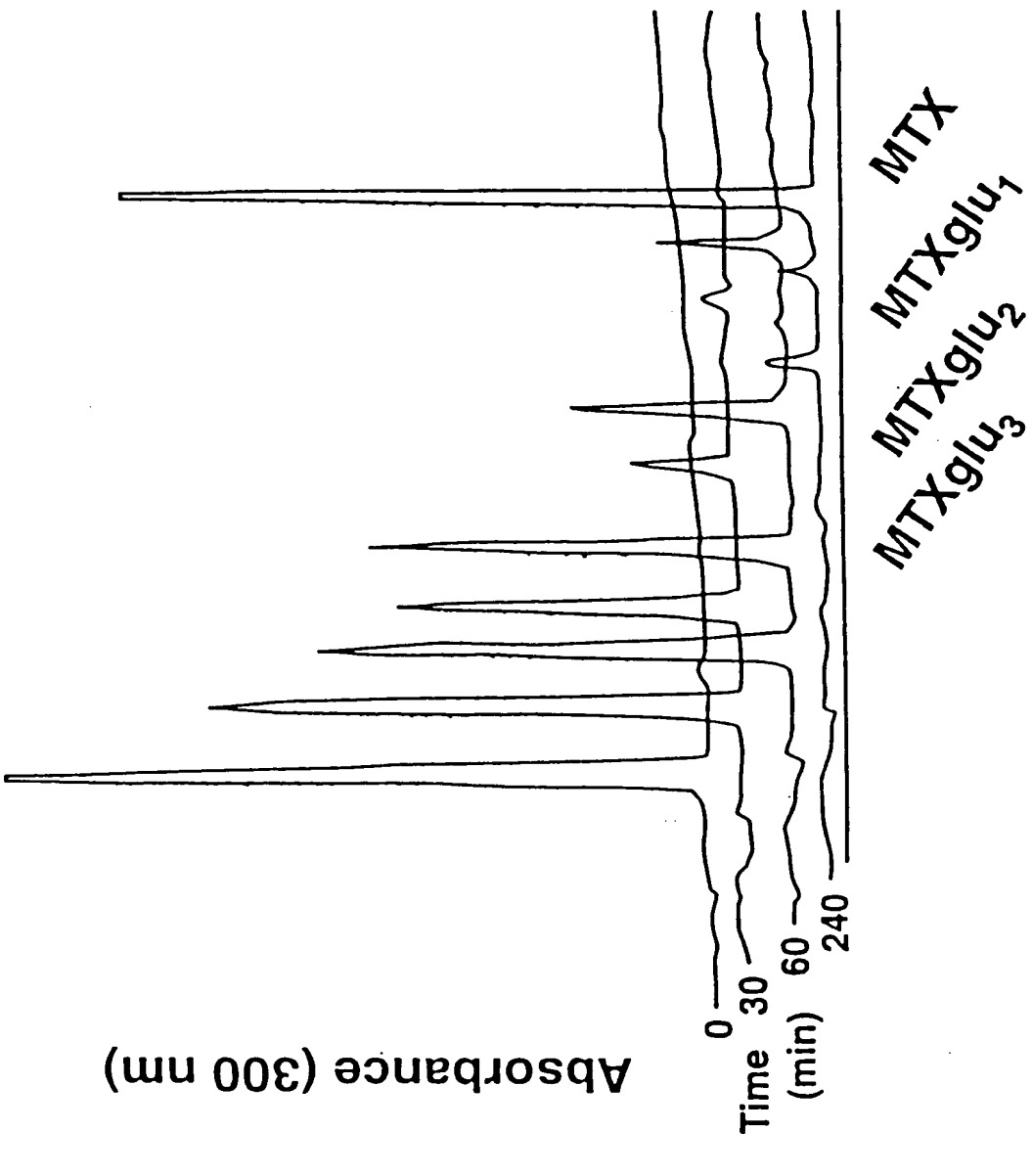
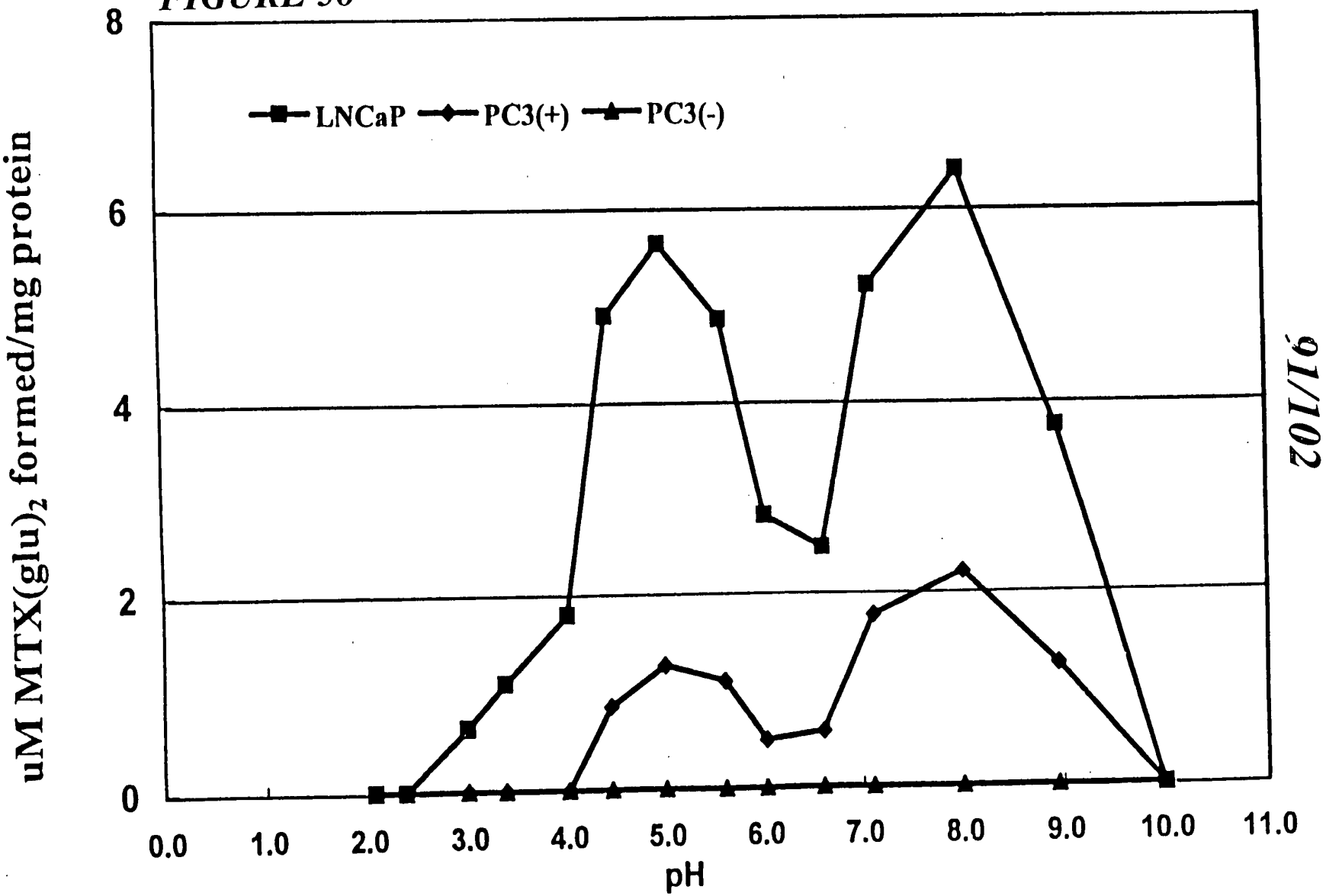

MTX
MTXglu1
MTXglu2
MTXglu3
0
30
60
240
Time (min)
Absorbance (300 nm)

FIGURE 57

FIGURE 58


LNCaP
PC3(+)
PC3(-)
uM MTX(glu)2 formed/mg protein
pH
0
2
4
6
8
0.0
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
10.0
11.0

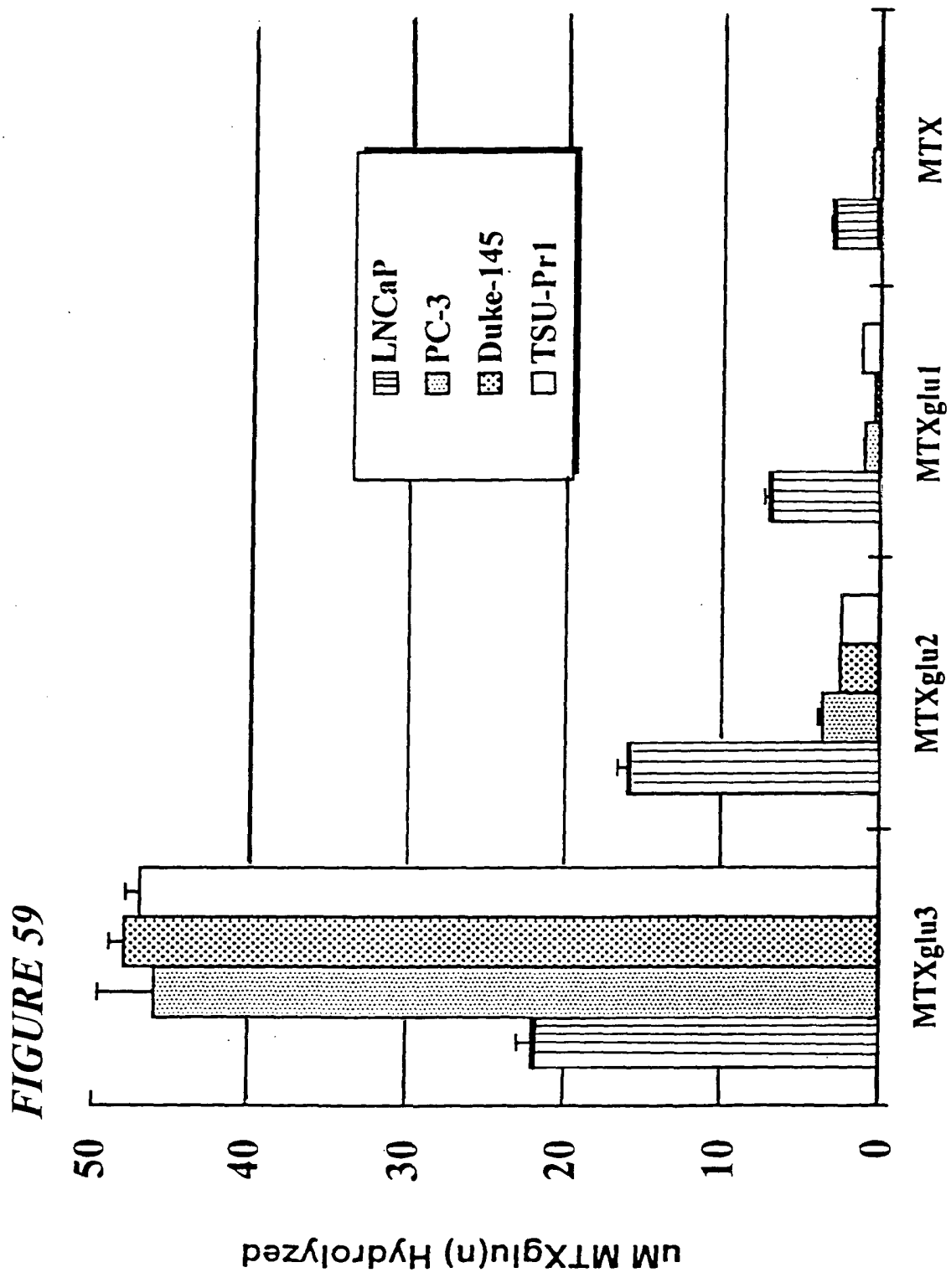

uM MTXglu(n) Hydrolyzed
50
40
30
20
10
0
LNCaP
PC-3
Duke-145
TSU-Pr1
MTXglu3
MTXglu2
MTXglu1
MTX

FIGURE 59

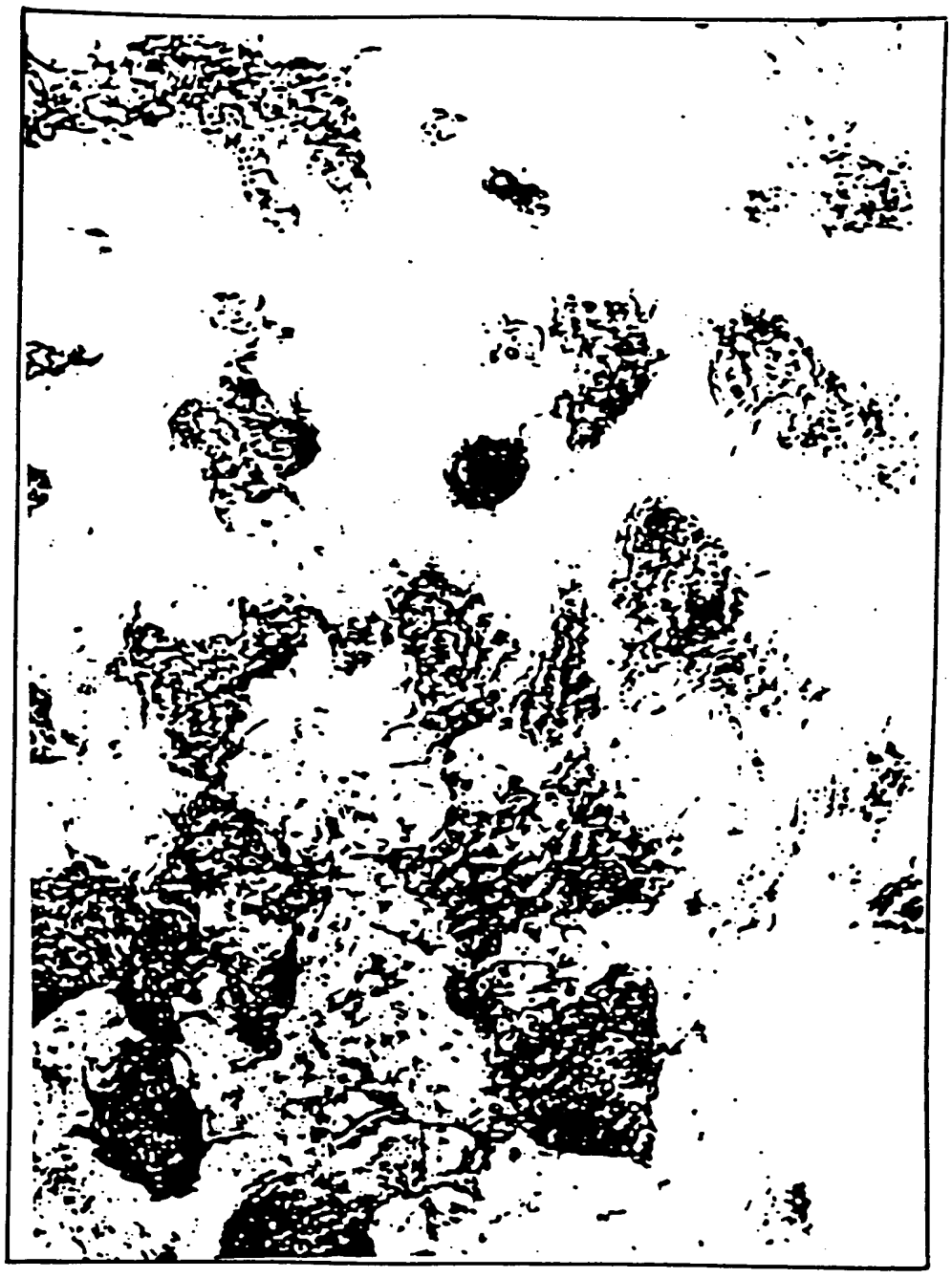


FIGURE 60A

FIGURE 60B

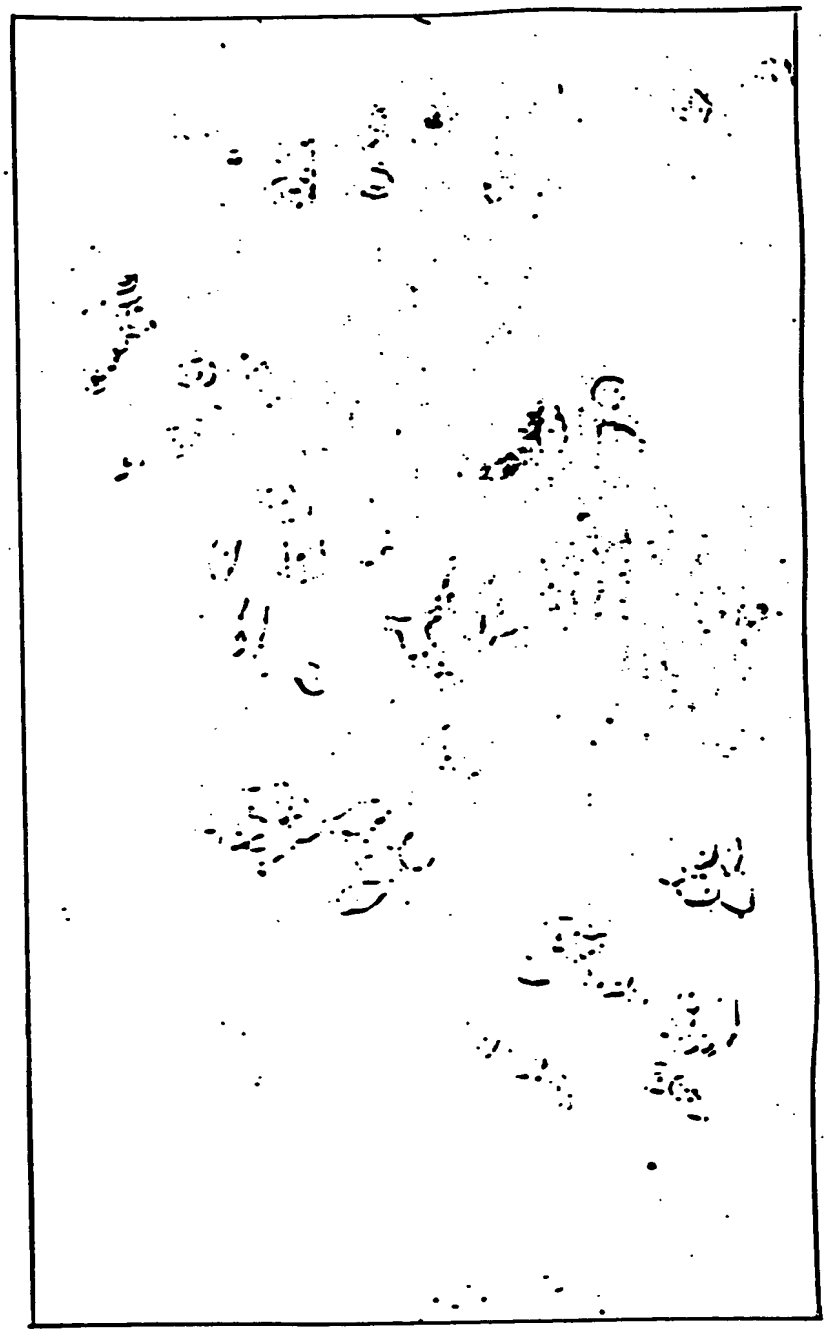


FIGURE 60C

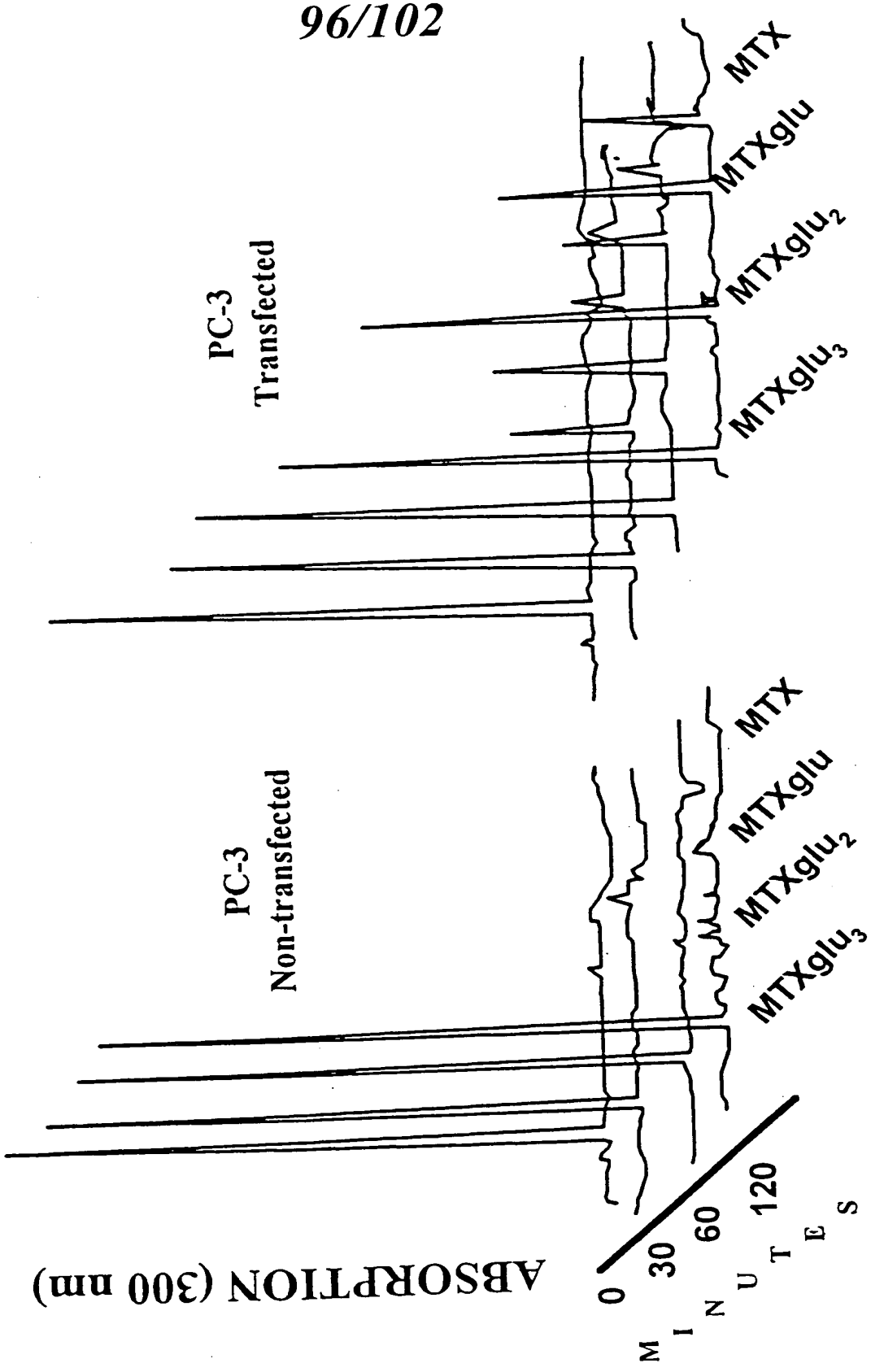
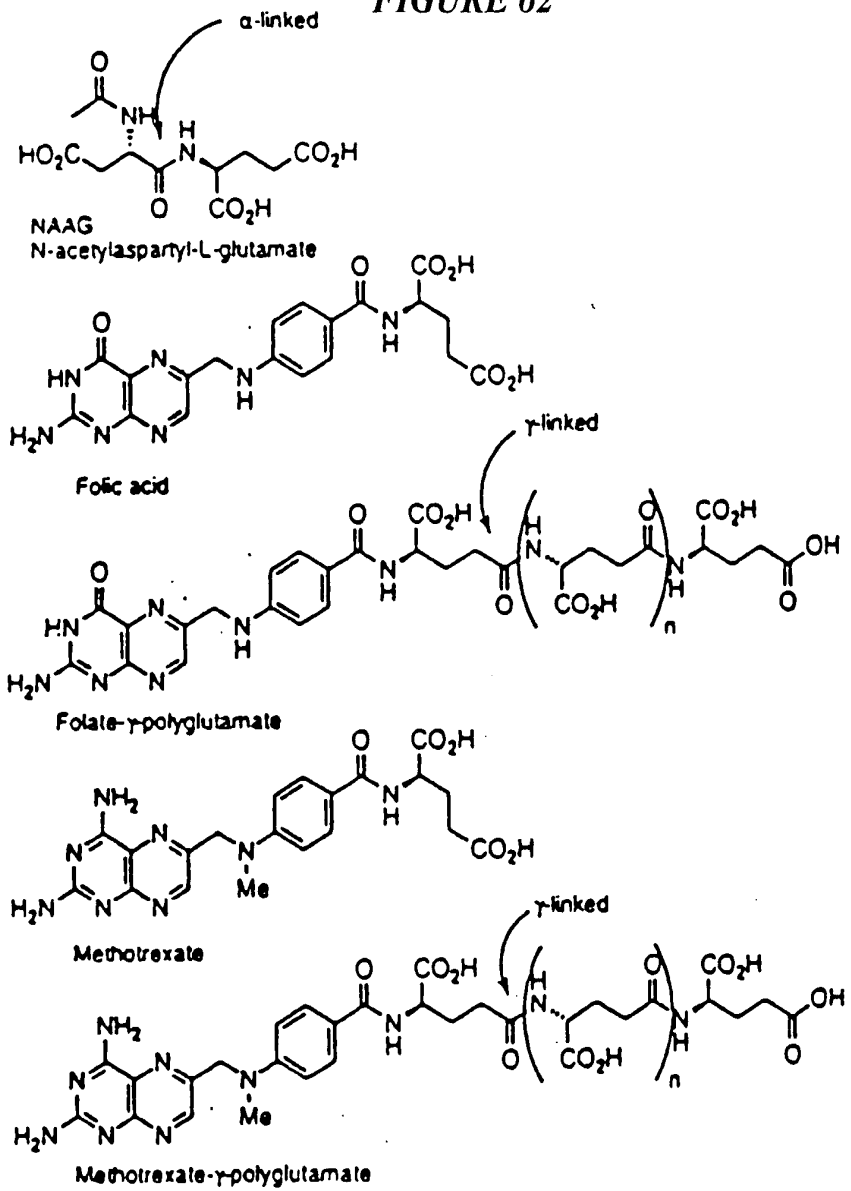

PC-3
Transfected
MTX
MTXglu
MTXglu2
MTXglu3
PC-3
Non-transfected
MTX
MTXglu
MTXglu2
MTXglu3
ABSORPTION (300 nm)
0
30
60
120
MINUTES

FIGURE 61

FIGURE 62


α-linked
NAAG
N-acetylaspartyl-L-glutamate
Folic acid
γ-linked
Folate-γ-polyglutamate
Methotrexate
γ-linked
Methotrexate-γ-polyglutamate

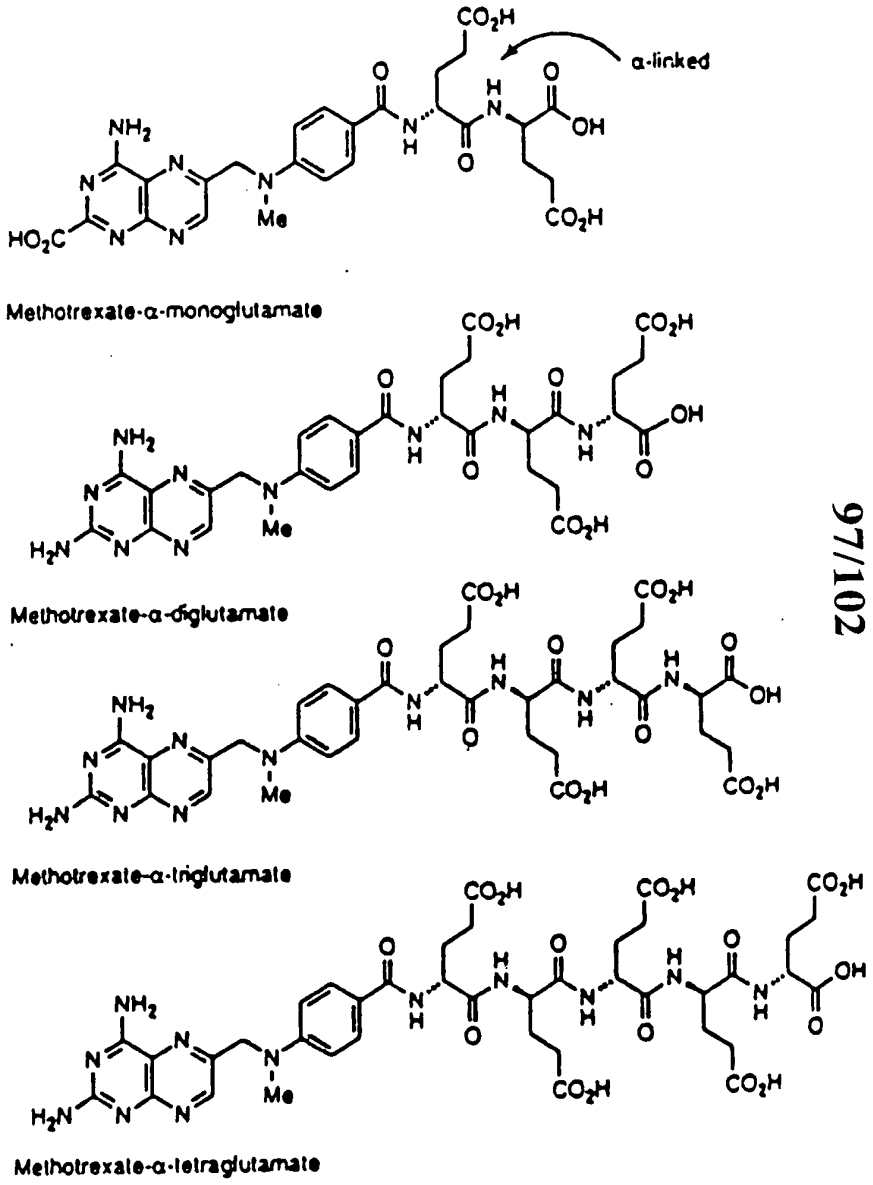

α-linked
Methotrexate-α-monoglutamate
Methotrexate-α-diglutamate
Methotrexate-α-triglutamate
Methotrexate-α-tetraglutamate

FIGURE 63A

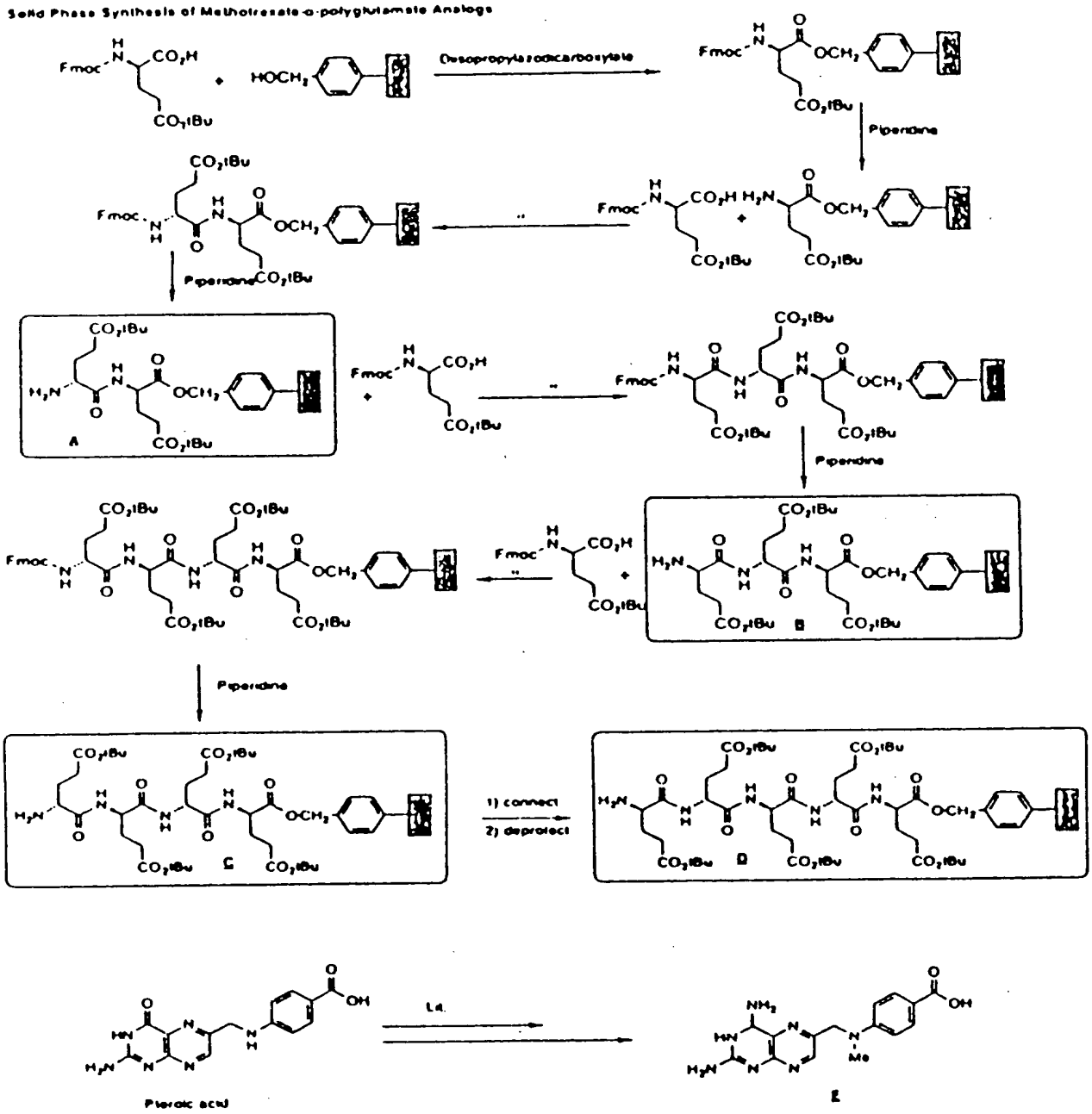

Solid Phase Synthesis of Methotrexate-α-polyglutamate Analogs
Diisopropylcarbodiimide
Piperidine
A
B
C
D
E
1) connect
2) deprotect
Pteroic acid

FIGURE 63B

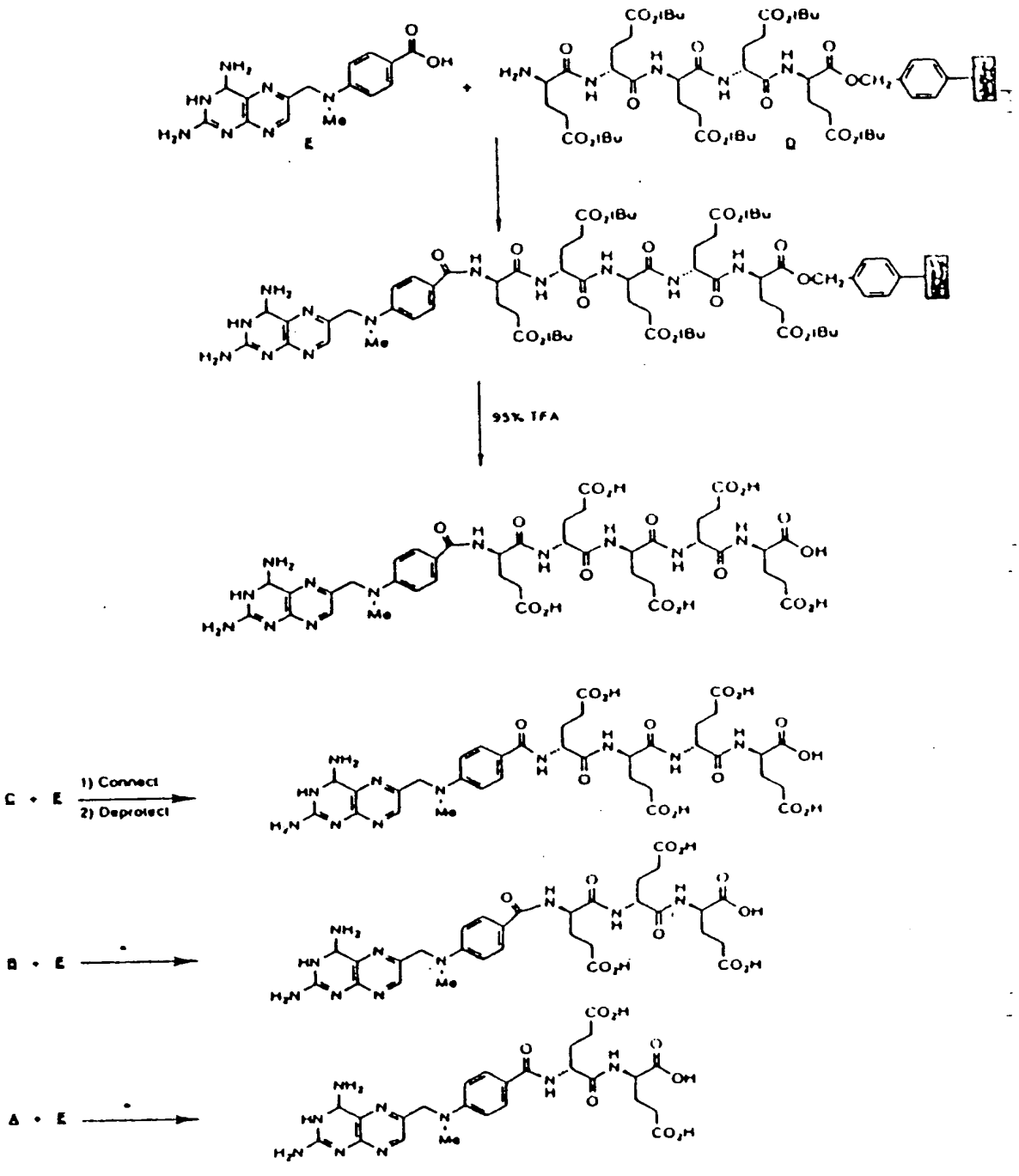


FIGURE 64

Sequence Analysis of microsatellite instability in PSM gene

Sample	Sequence	PSM EXPRESSION (IMMUNO STAIN)
Genomic	$T_9GC(TTTTG)_6(TTTG)_3T_7$	
LNCaP	$T_9GC(TTTTG)_6(TTTG)_3T_7$	positive
PC-3	$T_9GC(TTTTG)_8(TTTG)_3T_6$	negative
DU145	$T_{10}GC(TTTTG)_5(TTTG)_2T_7$	negative
T4 (tumor)	$T_{10}GC(TTTTG)_6(TTTG)_3T_7$	positive
N4(paired normal)	$T_9GC(TTTTG)_6(TTTG)_3T_7$	positive

FIGURE 65

Genomic Organization of PSM Gene

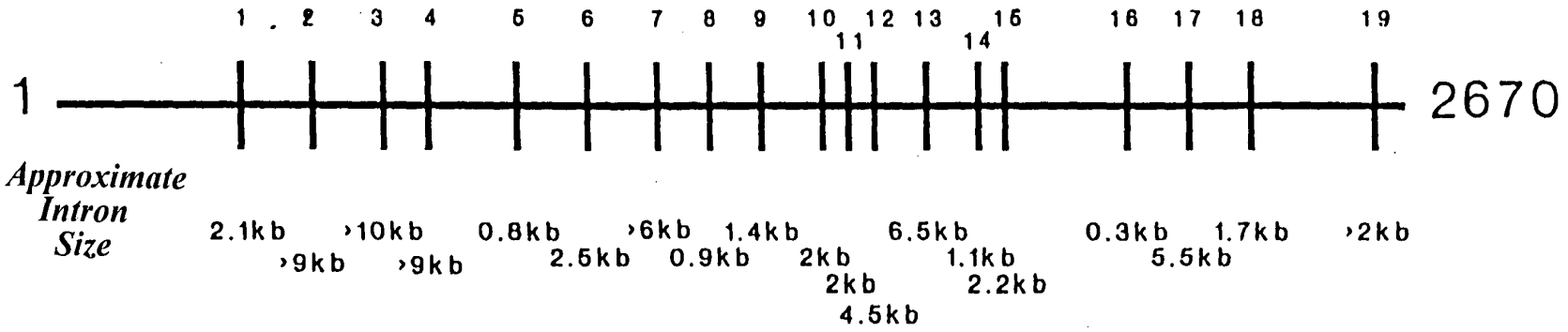


FIGURE 66

Location of microsatellite in PSM Gene